Homing in on getting out: Characterisation of SERA6, a putative malarial protease with a role in egress

Andrea Ruecker

March 2012

MRC National Institute for Medical Research
Division of Parasitology
Mill Hill, London
NW7 1AA

Division of Infection and Immunity
University College London

This thesis is submitted to University College London for the degree of Doctor of Philosophy

I, Andrea Ruecker, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Andrea Ruecker

Abstract

The human malaria parasite*, P. falciparum*, replicates within a membrane-bound intraerythrocytic parasitophorous vacuole (PV). The resulting daughter merozoites actively escape from the host cell in a process called egress. There is convincing evidence that proteases are key players in this step. These proteases could serve as excellent targets for the development of new antimalarial drugs. *P. falciparum* Serine Repeat Antigens (SERAs) form a family of 9 proteins all containing a central papain-like domain that identifies them as putative cysteine proteases. They are highly conserved throughout all *Plasmodium* species, and there is strong genetic evidence that they may play a role in egress. *P. falciparum* SERA6 is one of the most highly-expressed SERAs in asexual erythrocyte stages. In this study biochemical fractionation and indirect immunofluorescence analysis were used to confirm localisation of SERA6 to the PV. It was shown that SERA6 is a substrate for PfSUB1, a subtilisin-like protease which is crucial for egress and which is released into the PV just prior to egress. SERA6 is cleaved by PfSUB1 at 3 positions, releasing the papain-like domain. Processing of SERA6 by PfSUB1 is partially indispensable in *in vitro P. falciparum* parasites, as some mutations in SERA6 that block cleavage are not tolerated in the parasite. In addition, the putative catalytic Cys of SERA6 cannot be replaced with a non-catalytic Ala residue in the parasite, indicating SERA6 is an indispensable cysteine protease. Finally it was demonstrated that PfSUB1-mediated processing of the *P. berghei* orthologue of SERA6 converts it to an active cysteine protease. It can be proposed that SERA6 is a key player in a proteolytic cascade that leads to egress of the blood-stage malaria parasite.

Acknowledgements

First and foremost I would like to express my greatest gratitude to my supervisor Dr Mike Blackman. I regard him as an exceptional and unique scientist and person and I am most grateful to have had the opportunity to work with him during my PhD. I would like to thank him for his never ending patience and his deep believe in my skills.

Mike Shea, your wise words, unending guidance, support and sense of humour were simply invaluable. Without you the project would have not been possible. Many many thanks to each and every member (old and new) of the Blackman lab. Thanks for listening to my monologues that allowed me to perfect my rambling skills. I am convinced that without the support of the entire group I would have never finished this project. The support in and outside the lab has been irreplaceable to me.

Natalie, I could have literally not done my PhD without you. You are my porter carrying me up to the peak and still got never tired of listening to me although we never made it to the Kilimanjaro.

Fiona and Chris, I don't know what the lab would be without your exceptional organisational and social skills and I am so grateful for your guidance on every single day. Thanks for your patience and for making those long hours with all the parasites so much more fun. The lab would not work without the two of you. Robert thanks so much for your inspiration and the fact that you would never got tired to come up with new ideas and then sit down with me and helped me to realise them. Thanks Rob for your honest, cheerful and helpful advice throughout and especially the writing up phase. I would like to thank Sharon and Chrislaine for lightening up my days with your never ending knowledge and for your continuous and never ending patience towards my questions and discussions. Matt thanks for always listening to me and making me feel foolish in my mistakes.

I would like to thank all my fellow PhD students Caty, Maria and Sujaan for making the daily routine just that little bit more fun. I would also like to thank Malcolm for cheering me up in the early mornings. Thanks to the Holder lab especially Claire and Oniz for listening to the wanderings of my mind, thanks for supporting me, for your friendship and thanks for keeping up the fun. Thanks also to David, Judith and Ellen for your helpful discussion and frank advice.

I would also like to thank Volker Heussler, for providing us with an antibody which contributed hugely to the outcome of this PhD and all those who kindly gave me reagents for this work.

Thanks also to my old NIMR cottage house mates. Jen thanks for sharing the lows and the highs and for always supporting me every step along the way with your positiveness.

Meine Familie, ich kann euch nicht genug danken. Ihr habt immer an mich geglaubt, mir alles ermoeglicht und mir immer wieder gesagt, dass ich alles erreichen kann was ich moechte.

Finally thanks to James. It wasn't easy was it. You are my all. Ewig Dein, ewig mein, ewig uns.

List of Abbreviations

1

6xHis	Hexahistidine

A

Aa	Amino acid
Ab	Antibody
AMA1	Apical membrane antigen 1

B

BSA	Bovine serum albumin

C

CDPK	Calcium-dependent protein kinase
CN	Clear Native
CSP	Circumsporozoite protein

D

DAPI	4,6-diamidino-2-phenylindole
DBL	Duffy binding-like protein
DDT	Dichlorodiphenyltrichloroethane
dH_2O	Distilled water
ddH_2O	Double distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPAP3	Dipeptidyl peptidase 3
DTT	Dithiothreitol

E

E64	*trans*-Epoxysuccinyl-L-leucylamido(4-guanidino)butane, L-*trans*-3-Carboxyoxiran-2-carbonyl-L-leucylagmatine,N-(trans-Epoxysuccinyl)-L-leucine 4-guanidinobutylamide
EBA	Erythrocyte binding antigen
ECL	Enhanced Chemiluminescence

ECP1	Egress cysteine protease 1
EDTA	Ethylenediaminetetraacetic acid
EM	Immunoelectron microscopy
EPM	Erythrocyte plasma membrane
ER	Endoplasmatic reticulum
ETRAMP	Early transcribed membrane protein
G	
GDP	Gross domestic product
GST	Glutathione *S*-transferase
H	
H	Hour
HA	Haemagluttinin
HAP	Histoaspartic protease
HIV	Human immunodeficiency virus
HPLC	High pressure liquid chromatography
HRP	Horse Raddish Peroxidase
I	
IB	Inclusion bodies
IFA	Indirect Immunofluorescence Analysis
IP	Immunoprecipitation
iRBCs	Infected red blood cells
IRS	Indoor residual spraying
ITN	Insecticide treated bednets
K	
kDa	Kilodalton
M	
Mab	Monoclonal antibody
MAC	Membrane attack complex

MIC	Microneme protein
min	Minutes
MPP	Microneme protein protease
MSP	Merozoite surface protein
N	
NC membrane	Nitrocellulose membrane
P	
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PBST	PBS Tween 20
PCR	Polymerase chain reaction
PD	PfSUB1 prodomain
PEXEL	*Plasmodium* export element
PfAARP2	*P. falciparum* asparagine and aspartate rich protein 2
PfEMP1	*P. falciparum* erythrocyte membrane protein 1
PfRh	*P. falciparum* reticulocyte binding homologues
PfROM1	*P. falciparum* rhomboid 1
PfRhopH	*P. falciparum* rhoptry protein complex
PfSUB	*P. falciparum* subtilisin-like serine protease
PHMB	p-hydroxymercuribenzoate
PIC	Protease inhibitor cocktail
PKG	Protein kinase G
PoPS	Prediction of protease specificity
PPM	Parasite plasma membrane
PTRAMP	*Plasmodium* thrombospondin related apical merozoite protein
PV	Parasitophorous vacuole
PVDF	Polyvinylidene fluoride
PVM	Parasitophorous vacuole membrane
R	
RBC	Red blood cell

RESA	Ring-infected erythrocyte surface antigen
RON	Rhoptry neck
Rpm	rounds per minute
RT	Room temperature
S	
Sec	Seconds
SDS	Sodium dodecyl sulphate
SERA	Serine repeat antigen
T	
TB	*Mycobacterium tuberculosis*
TgMIC	*T. gondii* microneme protein
TPCK	N-tosyl-L-phenylalanine chloromethyl ketone
TX-100	Triton X-100
V	
v/v	volume/volume
W	
w/v	weight/volume
WHO	World Health Organization

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Malaria: The disease in all its aspects

1.1.1 A history of bad air

Malaria has been affecting humanity as a scourge for several thousand years and is thought to have killed more people than any other known human disease. It is one of the most devastating diseases of humankind, behind only tuberculosis (TB) and HIV. Despite one million deaths annually, its high profile cases of humans suffering from the disease known as “black fever” ranges from ancient Egyptian pharaohs, such as Tutankhamun (Hawass *et al.*, 2010;Timmann *et al.*, 2010) to modern world football players (Didier Drogba) or even pop stars (Cheryl Cole) and actors (George Clooney), keeping malaria in the news of non-endemic countries (other famous fatalities of malaria include Oliver Cromwell, Dante and Lord Byron). Even Christopher Columbus suffered from malaria on at least one occasion, indicating strongly that the disease has been affecting humans of all social classes throughout history and in recent days. Malaria has received many names throughout history, mainly derived from the symptoms sufferers were able to observe. The word malaria however derives from the Italian “mala (bad) aria (air)” as the disease was mainly associated with the fact that humans fell ill with malarial symptoms after sojourning in areas with strong odour of swamps or stagnant water. It was not before 1740 that the word “mal’aria” was introduced into the English culture by Horace Walpole. In the meantime it was portrayed as jungle fever, marsh fever, paludal fever or swamp fever. One name, black fever, was based on the colour of the urine passed by severely infected individuals.

Regardless of the attention malaria has received in more recent years, the origin of the malaria parasite, *Plasmodium spp.*, can be dated back much further. It is now generally thought that the human malaria parasite first occurred in our primate ancestors in Africa and evolved with human migration. A series of studies have attempted to date the origin of a variety of worldwide parasite strains by analysing mitochondrial DNA polymorphisms (Conway *et al.*, 2000;Joy *et al.*, 2003). It is suggested that recent parasite strains originate back as far as 50,000 years and that the common ancestors originated millions of years ago (Escalante *et al.*, 1994) with the most virulent human parasite, *Plasmodium falciparum*, diverging from its closest ancestor around 8 million years ago (Escalante *et al.*, 1995). However, contradictory reports suggest a rather younger history for the malaria parasite with a common ancestor being dated only around 5,000 to 60,000 years ago (Rich *et al.*, 1998;Volkman *et al.*, 2001;Conway *et al.*, 2002;Rich *et al.*, 2009). First descriptions of humans suffering malarial symptoms emerged in ancient Chinese medical reports around 2,700 BC, describing re-occurring tertian or quartan fevers (fever after 48 or 72 hours) with spleen enlargement and other symptoms (Ni, 1995). Hippocrates, however,

was the first consistent and epidemiologically detailed source to describe malarial symptoms in humans in early Greece (approximately 400 BC). In his medical diary “Of the Epidemics” Hippocrates describes symptoms as a “paroxysm” falling on “even days or odd days” and subsequently only re-occurring on even and odd days respectively. He also noticed “that if crises fall on days other than those mentioned above, there will be a relapse, and this may be deadly” and describes possible improvement “but it is essential to pay attention and know at which times the crises will lead to death and in which to recovery, or during which is there tendency to fair better or worse“ (Hippocrates, 2004). Before Hippocrates’s detailed first rendition of malarial symptoms, many more mentions of the human plague were found. Writings by Aristophanes, Aristotle, Sophocles and Plato (all around 400 BC) seem to describe what are now known to be malarial symptoms (Sherman, 1998). Even Homer (approximately 750 BC) poetically portrayed “Sirius, harbinger of fevers” caused by what he calls an “evil star which dominates the night sky at harvest time” most likely describing mosquitoes as the source of the fever which soldiers suffered from during the Trojan war (Homer *et al.*, 2007). Malaria of course was not only restricted to Europe and Africa, and reports of the disease were found globally with many fatalities. There is also experimental evidence for humans suffering from malaria in the ancient world. A first indication of the presence of malarial antigens in ancient mummies was found with a series of immunological tests (Miller *et al.*, 1994). To further confirm previous studies, mummy bone tissue samples were tested for the *pfcrt* region of *P. falciparum* in an elegant PCR based approach. Two out of 91 mummies from two different originating tombs in Egypt were identified to have suffered from malaria between 1,500 and 500 BC (Nerlich *et al.*, 2008). Of course there can be discussions about stability and longevity of such DNA samples; however there is no doubt that malaria has been playing a major role in the lives of many humans for thousands of years (Brier, 2004).

In 1882 Charles Louis Alphonse Laveran, a French military medical doctor, was the first person to unravel the connection between blood parasites and malaria. Laveran detected the parasites in the blood of Algerian soldiers. The disease had a colossal impact on the health of the French army and Laveran tirelessly attempted to uncover the secret cause of the disease. His observations published and distributed in scientific literature were at first dismissed by his medical colleagues. In 1884 he published his “Treatise on marsh fevers” but was only acknowledged with awards for his discovery in 1889. Laveran could not have established his findings and theories without previous work by Louis Pasteur. He undoubtedly prepared the ground for Laveran’s investigations by stating that most infectious diseases are caused by microbial germs in his “germ theory” in which microorganisms were proposed as the cause of many diseases (Pasteur, 1881). One of the major questions however still

remained unanswered until the year 1897 when Ronald Ross was the first investigator to uncover through the infections of birds that the malaria parasite was transmitted via the saliva of *Anopheles* mosquitoes (Ross, 1897).

Malaria has a very long history as scourge of humanity affecting people from all backgrounds and from all continents. However, it has not lost its relevance rather it has gained in importance as it causes devastation in the poorest of the poor countries, and affecting individuals that have no alternative but to live with the disease in silence.

1.1.2 Malaria: a socio-economic burden of the developing world

Malaria has not always been restricted to the developing world as up until the early 20th century it was widespread throughout the northern hemisphere, reaching its peak in the 19th century (Carter *et al.*, 2002). Mosquitoes and their breeding sites were long thought to be connected with the disease even before the first scientific evidence that an insect is in fact the transmission vector. Those breeding sites, which are mainly any areas with standing water and a temperature above a mean of 18°C, are found globally (Craig *et al.*, 1999). However, changing macro- and microclimates has caused eradication of the vector and the parasite in many countries and areas. Although the vector can still be found in some areas the parasite simply cannot survive in colder climates (Paaijmans *et al.*, 2009). Only a few non fatal autochthonous malaria cases have been located within those areas since eradication (Santa-Olalla Peralta *et al.*, 2010), although some of these cases are thought to be related to travel (Doby *et al.*, 1981). Malaria is currently thought to be responsible for around one million deaths annually worldwide with an incidence rate of clinical cases of approximately 300-500 million (Snow *et al.*, 2005;Guerra *et al.*, 2008). Due to the climate conditions required for malaria transmission the vast majority of these fatalities (80-90%), occur in sub-Saharan Africa where the incidence rate is as high as 212 million (see Figure 1), with 98% of all cases being caused by *P. falciparum* (WHO, 2009). Approximately 3.2 billion people live in 109 malaria-endemic countries, representing nearly half of the world´s population thus making malaria one of the most devastating diseases of our time, third only to HIV and TB (WHO, 2010). In recent years the impact of urbanisation on malaria mobility and morbidity in sub-Saharan Africa has been highlighted. It is suggested that the annual entomologic inoculation rates for malaria are reduced in urban settings compared to rural sub-Saharan Africa (Robert *et al.*, 2003); nevertheless, the estimated area covered by urban settings in Africa is only 1.6% of the total African surface (Keiser *et al.*, 2004). As few as 35 countries (30 in sub-Saharan Africa and five in Asia) account for around 95% of the global malaria death toll with children under the age of five and pregnant women being predominantly amongst those affected (WHO, 2003;Snow *et al.*, 2005;WHO, 2008;WHO, 2010). Of all the deaths due to malaria as

many as 85% occur in children under the age of five with every fifth childhood death being caused by malaria (Snow *et al.*, 2004). Those children living in endemic countries can expect to experience between one to six episodes of malarial infection before the age of five.

1.1.2.1 The economic impact of malaria

The countries with the highest malaria death toll are Nigeria, Democratic Republic of the Congo, Uganda, Ethiopia and the Democratic Republic of Tanzania. According to their per-capita gross domestic product (GDP) some of these countries are amongst the poorest in the world. Additionally they suffer from poor infrastructure, civil wars and/or political corruption resulting in financial deficits and a lack of investment from the private sector (Gallup *et al.*, 2001). The countries with the highest malaria prevalence and poverty are distributed identically over the world suggesting that malaria and poverty go hand in hand (Sachs *et al.*, 2002). Malaria accounts for a growth reduction of 1.3% in holo-endemic areas with up to 40% of public expenditure in these countries being spent on maintenance of health facilities and their staff, vector control and education and research required for managing the disease (WHO, 2009). The correlation between malaria prevalence and poor economic growth cannot be denied. Low GDP is an indicator for people being unable to afford necessary antimalarial treatment and prevention. These individuals will suffer from malaria and are incapable of attending work which in turn leads to growth reduction. Malaria has its most devastating impact on children, with prolonged and repeated periods of illness along with the many fatalities. In Kenya a child may miss approximately 15% of annual school days (Brooker *et al.*, 2000). In countries where full school education is the exception malaria has a devastating impact. Young children suffering from malaria need to be taken care of and carers will be unable to attend work, or if hospitalisation is required parents need to spend their little money on medication and care.

1.1.3 Apicomplexa

Malaria is caused by the obligate intracellular parasite *Plasmodium spp.* and is taxonomically ranked within the phylum Apicomplexa. To date over 5000 species can be found within the Apicomplexa and all are eukaryotic, protozoan, obligate, intracellular parasites of invertebrate and vertebrate hosts. The disease-causing parasites are not only a nuisance but also have a dramatic impact on human health, animal health and on the worldwide economy (Hung *et al.*, 2011). Some of the world's most devastating diseases are caused by members of the Apicomplexa. Antony van Leeuwenhoek was the first person to observe an apicomplexan parasite in 1674, identifying oocysts of *Eimeria stiedae* in the gall bladder of a rabbit. The first extensive

classification, however, was not attempted until 1987. Levine and colleagues (Levine, 1988) summarised all taxonomic knowledge about Apicomplexa which had a major impact on the understanding of the organisms. *Plasmodium* (malaria), *Toxoplasma* (toxoplasmosis), *Babesia* (babesiosis), *Cryptosporidium* (cryptosporidiosis), *Isospora* (isosporiasis), *Cyclospora* (cyclosporiasis) and *Sarcocystis* (sarcosporidiosis) are the seven genera which also infect and cause symptoms and disease in humans. An additional genus, *Theileria*, is an apicomplexan parasite with major impact on cattle livestock (Heussler *et al.*, 1999). *Plasmodium* and *Toxoplasma* however are of the utmost clinical and economic impact for humans, causing millions of deaths and huge financial losses each year (Jones *et al.*, 2001;WHO, 2010). *Toxoplasma* infects a third of the world's human population and although infection is generally asymptomatic, it can be fatal in immunosuppressed individuals, causing symptoms with co-infections such as HIV/AIDS (Kim *et al.*, 2008). *Babesia*, *Theileria* (both infecting cattle) and *Eimeria* (poultry) are major livestock pathogens and their impact on worldwide economies is a major challenge (Morrison *et al.*, 2006;Shirley *et al.*, 2007).

All apicomplexan protozoa share certain common features, whereas other features are only maintained by some of the genera. An apical complex, from whence the name Apicomplexa is derived, is composed of a group of organelles thought to be involved in penetrating or invading the host cell, and forms the most distinctive feature of the phylum. The apical complex is located at the apical end of the parasites and is comprised of a set of spirally arranged microtubules, the conoid, a cone-shaped cytoskeletal structure at the very apical end, a secretory body (the most prominent of which are the club-shaped rhoptries), polar rings and smaller electron dense secretory organelles (micronemes). The conoid is not essential for invasion across all the parasites, as it can only be found in the coccidian Apicomplexa, *Toxoplasma*, *Eimeria* and *Sarcocystis*. Coccidia have in common that at least one of the life stages requires to penetrate the intestinal epithelium, therefore a role for the conoid in that step has been suggested (Nichols *et al.*, 1987;Hu *et al.*, 2002);(Morrissette *et al.*, 2002). One additional essential structural feature is a chloroplast-like organelle, called the apicoplast, which is conserved throughout most of the Apicomplexa (Waller *et al.*, 2005). The apicoplast is a vestigial plastid of red algal origin (Janouskovec *et al.*, 2010) and was first found in malarial parasites of ducks (Kilejian, 1975). Its function has yet to be conclusively determined but many hypotheses, including plastid-like metabolic pathways, have been proposed (McFadden, 2010). All species across the Apicomplexa not only share common intracellular features but also have complex but similar life-cycles with sexual and asexual reproduction. A motile and infectious stage, the "zoite", which undertakes host cell invasion, is a common feature of the apicomplexan life-cycle. The host cell gliding form of the parasite injected by the mosquito vector

(Menard, 2001), named the sporozoite, actively invades the host cell by attachment and subsequent penetration. During schizogony it multiplies asexually and generates progeny within the invaded cell. Motile daughter cells are then released into the host and subsequently infect new cells. A few daughter cells will eventually further develop to form male and female gametes through gametogony. These combine to then produce new invasive forms during sporogony, the only sexual reproduction step in the life-cycle, allowing genetic recombination within species. Apicomplexan parasites carry a haploid genome, with one exception during sexual reproduction with the formation of a zygote (Morrissette *et al.*, 2002).

Toxoplasma gondii is the best understood member of the Apicomplexa and serves as a model organism for many other members including human pathogens, e.g. *Plasmodium spp*. or other pathogens of veterinary importance e.g. *Theilaria spp.*. The genome of *T. gondii* is amenable to manipulation in parasite cultures and allows for efficient stable and transient transfections of the parasite (Soldati *et al.*, 1993). Additionally, the phrase “size matters” can be applied literally to the understanding of the parasite’s biology. Most apicomplexan life stages are very small and test the limits of modern day microscopy. A *P. falciparum* infected red blood cell (iRBC) only measures 5 μM in diameter whereas *T. gondii’s* ability to infect a wide variety of cells can lead to infected host cells with diameters which are several fold higher than iRBCs.

1.1.4 *Plasmodium*: an introduction

Plasmodium spp. is a genus of obligate intracellular parasites with more than 200 species described to date. Reptiles (*reptilia*), birds (*avis*) and mammals (*mammalia*) are the main hosts for the disease (Cox, 2010). Morbidity and mortality caused in humans was the first impetus to investigate the origin of the disease. Interestingly the initial malaria parasites identified were not of human origin but *Plasmodium cynomolgi*, *Plasmodium inui*, and *Plasmodium pitheci* extracted from monkey blood samples (Garnham, 1966). Throughout the longest period of the 20th and the beginning of the 21st century four Plasmodium species were considered to widely infect humans: *Plasmodium ovale*, *Plasmodium vivax*, *Plasmodium malariae* and *P. falciparum* (Escalante *et al.*, 1995). In 2004 a fifth naturally acquired species, *Plasmodium knowlesi*, was shown to be responsible for fatalities in humans (Singh *et al.*, 2004) confirming Garnham’s hypothesis from 1957 that *P. knowlesi* might be another species causing human malaria (Garnham *et al.*, 1957). *P. knowlesi* was originally considered to only infect long-tailed macaques and human cases were thought to be occurring rarely until the realisation that it had potentially been misdiagnosed as *P. malariae* for many years based on their morphological similarities (Singh *et al.*, 2004;Cox-Singh *et al.*, 2008a;Cox-Singh *et al.*, 2008b;Lee *et al.*, 2009).

1.1.5 *Plasmodium* distribution and mortality

P. falciparum is the most virulent human malaria pathogen and is responsible for the vast majority of human deaths from malaria, especially in sub-Saharan Africa. It is transmitted by the female *Anopheles* mosquito. The intra-erythrocytic life-cycles, which cause most of the symptoms, between the five human malaria pathogens are surprisingly dissimilar. *P. knowlesi* undergoes its full intra-erythrocytic development within 24 hours, whereas *P. falciparum*, *P. vivax* and *P. ovale* require 48 hours for their full development. *P. malariae* however undergoes the longest intra-erythrocytic development, lasting 72 hours, and is known to establish an asymptomatic blood stage infection which can persist for decades. Although *P. vivax* and *P. ovale* are the only species known to develop dormant stages, called hypnozoites, within the human liver, which occasionally cause relapses months post-infection (Collins *et al.*, 2005;Coldren *et al.*, 2007;Imwong *et al.*, 2007). Not all of the five species are prevalent in all malaria endemic countries. *P. vivax* is widespread mainly in Asia and central and South America, affecting over 75 million people annually (Mendis *et al.*, 2001) and therefore coming second only to *P. falciparum* based on its epidemiology. The greater part of *P. ovale* derived infections can be seen in sub-Saharan Africa, mainly the western and eastern part, and some islands in the western pacific (Collins *et al.*, 2005). *P. malariae* distribution has been described as “patchy” but it is prevalent in most malaria endemic countries (Mueller *et al.*, 2007). *P. knowlesi* is almost exclusively spread around Southeast Asia (Lee *et al.*, 2011). The most virulent species *P. falciparum* is mainly found in sub-Saharan Africa where it accounts for more than 75% of the malaria cases (WHO, 2010). Therefore work in this study will focus on *P. falciparum*.

1.1.6 Experimental understanding of *P. falciparum*

P. falciparum has a haploid genome of 23 megabases distributed over 14 chromosomes and encoding approximately 5,300 genes. The genome is the most adenine and thymine-rich sequenced to date making *P. falciparum* relatively difficult to examine and explore with molecular genetic techniques (Menard, 2001;Gardner *et al.*, 2002). In 1975 Trager and Jansen were successful in culturing the pathogen in human blood *in vitro*, opening completely unexploited pathways in malarial research towards the understanding of the biology of the organism (Trager *et al.*, 1976). Despite the ability to explore *P. falciparum’s* biology *in vitro*, analysis of the entire life-cycle *in vitro* is a major challenge, since no *in vitro* model for mosquito-human interchange, or for the human immune response to malaria infections, has been established. Thus the search for an ideal system mimicking the human malaria pathogen keeps the research community on its toes. One goal was reached with the findings that *Plasmodium berghei*, a rodent parasite, can be used as a model for transmission from vector to the

mammalian host and for the first in-host stage with the sporozoite transit from the skin to the liver (Cao *et al.*, 2009b). Additionally, its close ancestry to *P. falciparum* makes it an intriguing model organism for further approaches.

1.1.7 Pathology: The disease in different aspects

P. falciparum is responsible for the most severe form of malaria and the intra-erythrocytic stages are causative of the associated pathology. Different severities of symptoms are associated with many factors and can range from mild and uncomplicated to severe malaria to cerebral malaria with a mortality rate of 19% (Newton *et al.*, 1998a). The clinical outcome of severe malaria is associated with many factors driving severe and fatal symptoms. Severe malaria in high transmission areas mainly affects young children between the age of six months and four years. Children up to the age of six months benefit from passive immunity transferred horizontally by maternal antibodies. After the first six months of their lives children need to rely on the establishment of their own immunity (see below) and mostly suffer from one to several malarial episodes until the age of four. By the time they reach the age of four, pathogenesis is rarely seen and generally has a mild outcome (Reyburn *et al.*, 2005). Adult sufferers can mainly be found in meso-endemic areas as they lack constant re-infection to maintain immunity against malaria. Clinical pathophysiological features of severe malaria are mainly characterised by metabolic acidosis, lactic acidosis (predominantly in infants), hypoglycaemia, severe anaemia, multi organ failure, respiratory distress, renal failure, hepatic dysfunction, microvascular obstruction and cerebral malaria. Uncomplicated malaria ranges from asymptomatic to flu-like symptoms, especially in immune individuals (Newton *et al.*, 1998b;Clark *et al.*, 2003;Francischetti *et al.*, 2008). High fever occurs every 48 hours coinciding with the rupture of infected red blood cells (iRBCs) and release of *P. falciparum* into the blood stream.

Cerebral malaria is caused by the adhesion of iRBCs to the microvasculature in the brain. This is possible due to parasite-induced RBC surface alterations. Adhesion of iRBCs in the vascular endothelium also leads to reduced tissue perfusion (Aikawa, 1988;Chakravorty *et al.*, 2008). By attachment of iRBCs to the lining of the vascular endothelium the parasite avoids clearance by the spleen, called sequestration. Sequestered parasites have been observed within various organs, for instance heart, lung, placenta and brain causing enlargement of the microvasculature. The key molecule involved in sequestration and in evading the host immune system is thought to be *P. falciparum* erythrocyte membrane protein 1(PfEMP1). PfEMP1 is parasite induced and highly expressed on the surface of iRBCs and interacts mainly with surface receptors of the microvasculature (Kyes *et al.*, 2001). PfEMP1 is encoded by

around 60 different gene copies in the parasite. Expressing one variant at a time enables the parasite to bind to tissue receptors and by switching approximately 2% per generation ensures evasion of the host's immune response (Roberts *et al.*, 1992;Bull *et al.*, 2005). Additionally *P. falciparum* causes rosetting, a process in which iRBCs adhere to non-infected RBCs. This leads to reduced deformability of non-infected RBCs, RBC lysis and obstruction to microvasculature causing additional pathology (Aikawa, 1988;Miller *et al.*, 2002a).

Although parasite factors, for instance cytoadherence, antigenic polymorphisms and variation are important, the role of host factors such as previous parasite exposure hence state of immunity or the host genetic background (sickle cell trait, thalassaemia), cannot be neglected.

1.1.8 Immunity and vaccination

Very little is known about immunity to *P. falciparum* derived malaria. It is understood however that different pathology in humans is influenced by both parasite and host factors. The differences between severe malaria, mild malaria and the establishment of asymptomatic infections are thought to occur based on previous exposure to the pathogen. Two types of immunity can be described: anti-disease, which mainly alleviates the symptoms but does not affect the parasite, and anti-parasite, which is involved in reducing parasitaemia. Humans living in holo-endemic areas with constant re-infections have a level of anti-disease immunity but not sterile immunity (Kyes *et al.*, 2007). It was shown that constant re-exposure to *P. falciparum* is required to achieve a state of premunition. People from holo-endemic areas that moved to non-endemic countries suffer from malaria if being infected again after living away for a certain amount of time.

As previously mentioned newborn children up to six months of age obtain low level immunity through maternal antibodies. However, between the age of six months and approximately four years children need to establish their own protective immunity. They will acquire some immunity only if exposed to the parasite, but at the same time they might suffer from severe malaria, cerebral malaria or possibly leading to death (Snow *et al.*, 2001). Acquiring immunity to malaria is very effectively challenged by the complex life-cycle of the parasite. The parasite is exposed directly to the human immune system for only a brief amount of time during its life-cycle - first, as pre-erythrocytic infective sporozoites before invading hepatocytes and second as merozoites are released into the host's blood stream before invasion. Immunity is mainly established against parasite-derived antigens which are expressed on the host cell surface. Even though immunity to a repertoire of these variant surface proteins may be slowly acquired, variant erythrocytic surface antigen expression (antigenic

variation) attenuates the possibilities of inducing a naturally acquired immune response against these variants, and, more importantly, affects host immune memory (Bull *et al.*, 2005). The host's genetic pre-disposition also contributes to pathology and some RBC polymorphisms are understood to lead to milder pathology (Wellems *et al.*, 2009). Individuals heterozygous for sickle cell anaemia, named sickle cell trait (HbAS), suffer milder malaria pathophysiology and lower parasitaemia compared to healthy individuals (Allison, 1954). The exact mechanisms are unknown. It is believed that parasites are less capable of invading the crescent shaped erythrocytes and infected erythrocytes undergo reduced sequestration limiting parasite survival hence restricting pathology. Additionally it has been suggested that HbAS causes the infection to persist longer in the body, therefore a better immune response can be established (other theories for the HbAS effect have been proposed; Williams. A third of all sub-Saharan Africans are heterozygous for sickle cell anaemia which increases fitness of affected individuals towards malaria, suggesting selection for sickle trait sufferers has taken place in endemic areas. Other genetic alterations that improve fitness against malaria, are α-thalassaemia, causing a reduction in severe malaria and adhesion of infected and non-infected RBCs and glucose-6-phosphate dehydrogenase (G6PD) deficiency, leading to increased phagocytosis of early stage parasitised RBCs (Cappadoro *et al.*, 1998;Lopez *et al.*, 2010).

Although the *P. falciparum* induced immune response still raises many questions, attempts to develop effective and protective vaccines to induce anti-disease immunity are continuously on their way (Crompton *et al.*, 2010). However, these efforts are hindered by the complex life-cycle of the parasite, antigenic variation and the above mentioned lack of understanding of the pathogen-host interaction in the immune response. PfEMP1 plays a central role in host-parasite immune development and is exclusively expressed by transcription of *var* genes (Gardner *et al.*, 2002), making it a rather less interesting target for the development of vaccines because of antigenic variation. A few non-antigenic variation encoded transcripts provide targets for the immune response and are considered vaccine candidates (Riley *et al.*, 2006). The circumsporozoite protein (CSP) on the surface of infective sporozoites can be targeted by the immune response after the inoculation of irradiated or genetically attenuated sporozoites (Kumar *et al.*, 2006). Protection is believed to be most efficient if sporozoites invade hepatocytes but cannot develop completely. Additionally a vaccine carrying a part of CSP in fusion with a surface coat protein of Hepatitis B virus, called RTS,S, has succeeded as far as clinical Phase III trials, with conferring protection with efficacies of between 40% and 50% in Phase II (Alonso *et al.*, 2005;Casares *et al.*, 2010). The above mentioned vaccination approaches target pre-erythrocytic stages,

preventing the establishment of blood stage infections by diminishing intra-hepatocytic development.

Antibodies produced against blood stage molecules, e.g. the merozoite surface protein (MSP1) expressed on erythrocytic stage merozoites, reduce parasite growth by inhibiting RBC invasion. Development of vaccines based on recombinant MSP1 is ongoing (Christian Epp, unpublished; Ramasamy. Vaccines against blood stage molecules however suffer widely from low-level efficacy.

1.1.9 Malaria control

Attempting to control malaria has been a challenging battle for the last few centuries. Many approaches were taken to eradicate the parasite and the vector, or if not to eliminate then to invent strategies for coping with the disease. Although mortality appears to be declining (WHO, 2008;WHO, 2010), malaria still is as prominent as it has been for many centuries. The struggle to understand all the key parts in the parasite's life-cycle, the host-parasite interaction, and the resulting immune response have prepared the ground for new approaches in the antimalarial battle. Efforts were additionally turned towards vector elimination, especially with the new availability of chemicals attacking *Anopheles* mosquitoes.

Over the past decades both organisms have evolved high levels of resistance to insecticides (Ranson *et al.*, 2000;Catteruccia, 2007) or antimalarial drugs (Takebe *et al.*, 2007). Therefore numerous obstacles still have to be overcome in controlling the disease. However, malaria control has not always appeared that way. In the 1950's malaria eradication was believed to be imminent (Russell, 1955;Hay *et al.*, 2004). In recent times the targets have been re-evaluated and the short term main objectives are malaria control independently from eradication (Moonen *et al.*, 2010). Improving malaria control means improving monitoring malaria morbidity as misdiagnosis leads to unnecessary treatment causing parasite resistance against antimalarial drugs.

The fight against malaria can be addressed from diverse angles; reducing transmission and infection and administering disease management.

1.1.9.1 Parasite control: Antimalarial drugs

Drugs with antimalarial capacity are currently the best therapy for decreasing morbidity and mortality. Nearly all antimalarials are developed to target the most severe form of malaria caused by *P. falciparum* by attempting to kill the parasite within infected erythrocytes. The search for effective and suitable medication has been ongoing.

Ever since malaria has been affecting humans people relied on naturally derived compounds primarily from plants or extracts thereof. The anti-fever properties of a natural compound qinghao, *Artemisia annua* (sweet wormwood), were first

described in China in 168 BC. It was frequently used to treat fever and is considered to have been in use widely long before that time (Klayman, 1985). The most effective recent antimalarials are semi-synthetic artemisinin derivatives and their analogues and these are now the recommended first-line treatment. They target the intra-erythrocytic stages of the parasite, reducing parasitaemia and disease symptoms. They are used in severe malaria mainly in combination therapy to avoid recrudescence and parasite resistance (Maude *et al.*, 2010). Two of the most widely used applied semi-synthetic artemisinin analogues are artemether and artesunate. Unfortunately, the first resistant parasite strains have now been reported (Saralamba *et al.*, 2011). Despite its long history, it took until the early 1970's for researchers to gain interest into artemisinin, coinciding with the first severe cases of *P. falciparum* resistance to another group of antimalarials. Chloroquine was discovered in 1934 as a methylene blue derivative and first administered in the 1940's (Coatney, 1963). The mode of action is conferred by chloroquine being passively taken up into the food vacuole of the parasite (residing within an erythrocyte) where chloroquine undergoes protonation, accumulates within the food vacuole and inhibits the sequestration of hem in the form of haemozoin. This causes the parasites to become slowly intoxicated and die. Combined with vector control, chloroquine was one key tool in the WHO malaria eradication programme launched in 1955 (WHO, 1999). Unfortunately, resistance emerged quickly, mostly associated with mutations within the *P. falciparum* chloroquine transport channel (PfCRT) in the food vacuole (FV) membrane. These mutations are thought to allow the protonated chloroquine to exit, leaving a resistant *P. falciparum* (Payne, 1987;Fidock *et al.*, 2000;Wellems *et al.*, 2001). Chloroquine resistance had a disastrous effect on morbidity and mortality caused by *P. falciparum* and *P. vivax* and led to termination of the WHO malaria eradication programme in 1972. Resistance emerged as a severe effect of over usage and large scale administration without effective control strategies in place (all individuals suffering from malaria-like symptoms were treated without confirmed diagnosis). Resistance is wide-spread, and additionally it can be detected in travellers returning from holo-endemic countries to the UK; see Figure 2). Because of its cost-effectiveness chloroquine is still used in many places as a first-line treatment. Other antimalarials are for instance administered as a combined therapy between atovaquone and proguanil inhibiting the parasite's dehydrofolate reductase (DHFR), e.g. Malarone™. Mefloquine, a quinine analogue is also used, but with resistances emerging as quickly as one year after administration (Wongsrichanalai *et al.*, 2002;Price *et al.*, 2004;Sutherland *et al.*, 2008).

Many effective antimalarials mimic natural compounds or are in fact derived from direct plant extracts. Intelligent drug design has only come into place within the last decades and understanding the parasite's biology is of utmost importance. Until

then re-design of old drugs, rediscovery of old molecules, drug combinations and the application of natural compounds are required to not elevate the malaria death toll.

1.1.9.2 Vector control and disease prevention

In efforts to reduce malaria prevalence all possible steps must be taken. Therefore vector control can play as important a role as parasite control. Vector management can be approached from diverse angles as the mosquito undergoes two morphologically distinct phases, a larval and adult stage. Prevention of breeding or killing the larval and the adult forms would interfere with transmission as would the prevention of the blood meal by adult female mosquitoes. Different vector control measures include the distribution of insecticide-treated bed nets (ITNs) to reduce bite count, varying methods of insecticide spraying to attack either the adult mosquito or breeding sites, and finally the latest approaches to establish transgenic mosquitoes. Attacking adult mosquitoes has many advantages over larvae control. Larval breeding sites are not only extremely widespread but also inaccessible. Aquatic breeding sites require frequent treatment with insecticides whereas ITNs or solid indoor wall spraying persist for up to six months. Additionally the maturation of the eggs within the female mosquito takes up to 12 days during which time it can return up to three times for a blood meal, increasing the likelihood of elimination. In the past the main focus was indoor residual spraying (IRS) or the large scale outdoor spraying of insecticides with dichlorodiphenyltrichloroethane (DDT). In combination with chloroquine treatment of infected individuals excessive DDT spraying was used for the malaria eradication programme in the 1950's. In some low transmission areas the programme had a colossal impact on the reduction of malaria prevalence (WHO, 1999). However, little impact was achieved in high transmission areas where health concerns about the extremely widespread and extensive application of DDT were raised as quickly as reports of resistant mosquitoes, which lead to abortion of insecticide application and to the arrest of the entire eradication programme (Wellems *et al.*, 2009). The focus has now returned to IRS with pyrethroids as the chemical of choice as the vectors are known to feed and rest indoors making walls covered in insecticides an applicable and controllable approach. Furthermore bed nets have been useful in controlling malaria transmission since ancient times. Even Cleopatra was thought to have used bed nets to keep the biting nuisances away at night. Distribution of ITNs is easier to monitor than the use of chemicals and they are cost-effective. A considerable amount of effort has been combined between several non-profit organisations and profit based companies operating in some African countries. For example lorries which are required for the transport of coffee beans from place of cultivation to distribution centres can also serve to transport ITNs (personal communication, All Party Parliamentary Meeting for Malaria

and Neglected Tropical Diseases 2010). 40-60 million ITNs were distributed in 2008 and in some African countries the household ITN ownership was up to 80% (WHO, 2009). However, the data can be misleading as some countries, e.g. Nigeria, have less than 10% household ITN ownership.

Although there has been a relative long history in the application of sterile mosquitoes (Knipling, 1959) it was not before the completion of the *A. gambiae* genome project in the past decade that practical advances were made (Holt *et al.*, 2002). Progress has been made towards the establishment of genetically driven systems as a basis of transgenic vector control. Population replacement and population eradication are considered as two main strategies in vector control. An alternative approach has been taken with the development of sterile mosquitoes to reduce or eliminate *Plasmodium* transmission (Thailayil *et al.*, 2011). Methods like the use of symbionts for instance *Wolbachia* and the direct offensive against *Plasmodium* in the host´s midgut and salivary glands to prevent transmission are now in use (Christophides, 2005;Hughes *et al.*, 2011;Isaacs *et al.*, 2011).

1.2 The *P. falciparum* life-cycle

P. falciparum is transmitted by mosquitoes of the genus *Anopheles* and exhibits a complicated life-cycle of asexual multiplication in the human host and sexual reproduction in the dipteran vector (see Figure 3). Of the 400 *Anopheles* species distributed worldwide, only 30 are considered to be of major importance in malaria transmission (Tuteja, 2007), with *Anopheles gambiae* being the most important in transmission of human malaria in Africa (Holt *et al.*, 2002). Even though the parasite causes sequelae and morbidity mostly in the human host (Dawes *et al.*, 2009), man acts merely as an intermediate stage, and the parasite´s sexual reproduction is completed exclusively in the insect vector. Only female mosquitoes transmit malaria as male do not feed on human blood, as the blood meal is required for the full development of the eggs within the female mosquito.

Throughout its life-cycle *P. falciparum* varies between motile, invasive, encysted and intracellular sexual and asexual stages. Other *Plasmodium* species additionally form dormant forms. Infective sporozoites are inoculated from the female mosquito`s salivary glands into the subcutaneous tissue (and, less frequently, directly into the blood stream) via injection by the proboscis of the female mosquito during the blood meal (Daily, 2006). Sporozoites breach the blood vessels of the skin, migrate via the blood stream to the liver and actively invade hepatocytes. For reasons that are poorly understood sporozoites traverse several hepatocytes before the final invasion of a cell where they undergo exo-erythrocytic mitotic division (exo-erythrocytic schizogony (Prudencio *et al.*, 2007;Sturm *et al.*, 2007). Eight to ten days after invasion, hepatic

schizonts rupture, each releasing several thousand merozoites into the host circulation in membrane-surrounded vesicles called merosomes (Sherman *et al.*, 2004;Sturm *et al.*, 2006). Once in the blood stream, infective merozoites actively invade RBCs, beginning the blood stage cycle (intra-erythrocytic cycle). Post-invasion, the parasite, which resides in a parasitophorous vacuole (PV) with its own parasitophorous vacuolar membrane (PVM) develops into the ring stage (0-24 h post invasion), followed by the trophozoite stage (24-36 h post invasion) and finally the schizont stage (36-48 h post invasion). Throughout the intra-erythrocytic the parasite acquires its nutrients in a very sophisticated way by the breakdown of the host cell haemoglobin and by the uptake of nutrients from its surroundings (Tilley *et al.*, 2007;Polonais *et al.*, 2010). Erythrocytic schizogony coincides with multiple rounds of nuclear division of the trophozoite to produce mature schizonts that then undergo cytokinesis to finalise asexual reproduction and produce numerous daughter cells (merozoites). The host cell finally ruptures in an active process called egress and 16 to 32 merozoites (in the case of *P. falciparum*) are released into the blood stream to infect fresh RBCs. Not all released merozoites undergo another erythrocytic cycle. Some will differentiate, may be triggered by the level of parasitaemia and the availability of nutrients (Nacher *et al.*, 2002), into sexual stages called macrogametocytes and microgametocytes that mature but are incapable of reproduction in the human host. Their ingestion by a feeding *Anopheles* activates further maturation, such as exflagellation of microgametocytes (Billker *et al.*, 2004). Altered pH and temperature conditions, xanthurenic acid and calcium signalling via a cGMP-dependent protein kinase are some of the known factors that facilitate DNA replication and cell division and induce differentiation of the male and female gametocytes to gametes (Vlachou *et al.*, 2006;McRobert *et al.*, 2008). Soon after the gametes fuse, the resulting elongated ookinete penetrates the mosquito´s midgut epithelium and matures to an oocyst while embedded in the midgut epithelium. Finally, sporozoites which develop in the oocyst breach the oocyst wall and migrate to the salivary glands of the dipteran host (Ghosh *et al.*, 2009) ready to be injected into a new human host.

Although *P. falciparum* survives within a rather multifaceted life-cycle with varying developmental stages and a range of host cells it is important to mention that despite all complexity conservation within various life-cycle steps can be found. Active invasion and active escape are a common theme for *P. falciparum* and can be found in all hosts throughout all life stages.

1.3 Intra-erythrocytic stage

The complexity of the *P. falciparum* life-cycle is a major obstacle that can only be overcome by completely understanding all the individual events. Despite involving

various different developmental stages, one in particular is the most problematic for humans. The intra-erythrocytic (see Figure 4) stage starts after the release of several thousand merozoites from hepatocytes via merosomes into the host's blood stream.

1.3.1 Getting in: Invasion

Invasion by *Plasmodium spp.* is a rapid and tightly regulated process that was first characterised and is best understood in *P. knowlesi* although it is now believed that *P. falciparum* undergoes a very similar process (Bannister *et al.*, 1975;Dvorak *et al.*, 1975;Aikawa *et al.*, 1978;Bannister *et al.*, 1990). Merozoite invasion of RBCs is a necessity to ensure survival of the parasite and it is the step initiating pathology in the human host. Successful invasion begins after the initial weak and reversible attachment between the merozoite surface and the virgin erythrocyte (Bannister *et al.*, 1990). This step is required for recognition and identification of the host cell as only RBCs can be invaded at this stage. Subsequent re-orientation wherein the apical part of the merozoite is in contact with the RBC surface is required before the parasite can invade the RBC. Re-orientation coincides with the specific formation of a tight junction orchestrated by the release of proteins from parasitic secretory organelles onto the RBC surface whereupon the former weak attachment becomes irreversible (Bannister *et al.*, 2003). To now enter the cell the tight junction moves towards the posterior end of the merozoite driven by a parasite actin-myosin motor (Keeley *et al.*, 2004;Baum *et al.*, 2006).

Understanding the molecular basis of invasion requires some knowledge of the merozoite's ultrastructure and its organelles (see Figure 5). Merozoites are membrane bound invasive stages that assemble after several asexual nuclear divisions of a parent very late in the parasite's intra-erythrocytic life-cycle. The surface of the merozoite is coated with glycosylphosphatidylinositol (GPI) anchored, peripheral and transmembrane domain containing proteins that are discharged from merozoite secretory organelles before or during and after egress and invasion. They are generally called merozoite surface proteins (MSPs) and the most prominent member covering the majority of the surface is MSP1 and is the most studied MSP. It is essential for parasite survival, is a major vaccine candidate (see 1.1.8) and is thought to initiate the first weak attachment between merozoite and the RBC surface (Holder *et al.*, 1982;Holder *et al.*, 1992;O'Donnell *et al.*, 2000). Band 3 and heparin-like molecules are considered possible RBC surface ligands to interact with MSP1 (Goel *et al.*, 2003;Boyle *et al.*, 2010a). Antibodies against MSP1 abolish parasite invasion *in vitro* and protect against blood stage challenge *in vivo* (Holder, 2009;Woehlbier *et al.*, 2010). MSP1 forms a merozoite surface complex with the peripherally associated family members MSP6/7 (Stafford *et al.*, 1994;Stafford *et al.*, 1996) and regulated extensive

maturation of the complex prior and throughout invasion is thought to be required for parasite survival (Pachebat *et al.*, 2007;Child *et al.*, 2010). Different MSPs are believed to hold different roles in invasion as they are unevenly distributed across the merozoite surface (Sanders *et al.*, 2005). Further GPI-anchored surface proteins are Pf12, Pf38, Pf92 and the putative Pf113 with some of them being refractory to genetic disruption in the parasite suggesting an important role (Cowman *et al.*, 2006).

1.3.1.1 Irreversible attachment and tight junction formation

After the initial weak attachment (see Figure 6A) with the RBC ligands the merozoite re-orientates for an irreversible contact between the apical end and the host cell surface, positioning the secretory organelles juxtaposed to the RBC surface (see Figure 6B). This process is thought to be triggered by the accumulation of proteins towards the apical end of the merozoite possibly coinciding with discharge from micronemes of proteins that then translocate across the merozoite surface (Lew *et al.*, 2007). Apical membrane antigen 1 (AMA1) is believed to be translocated from micronemes across the merozoite surface prior to invasion but probably does not display any role during the initial weak attachment (Mitchell *et al.*, 2004). AMA1 is highly conserved throughout Apicomplexa and is essential for parasite invasion in *P. falciparum* and *T. gondii* (Triglia *et al.*, 2000;Silvie *et al.*, 2004;Mital *et al.*, 2005).

Now being correctly positioned the merozoite apical organelles continue to secret a subset of proteins to promote invasion and to form an irreversible tight junction between the merozoite and the host cell (see Figure 6C/D). The discharge from organelles is strongly believed to be tightly regulated by calcium efflux through signal transduction cascades, as intracellular calcium chelators inhibit secretion, but secretion can be induced with calcium ionophores in *T. gondii* (Carruthers *et al.*, 1999a;Carruthers *et al.*, 1999b). Additionally, specific inhibition of *T. gondii* calmodulin-like domain protein kinase 1 (CDPK1) prevent attachment and micronemal secretion (Kieschnick *et al.*, 2001).

Candidates for the establishment of the irreversible interaction (tight junction) between the RBC and the apical end of the merozoite are proteins of the Duffy binding ligand (DBL) family and *P. falciparum* reticulocyte binding homologues (PfRh). DBLs, which are characterised by a cysteine-rich domain towards the N-terminus, were first identified in *P. vivax* and have been shown to interact with the Duffy antigen on the surface of RBCs (Wertheimer *et al.*, 1989;Fang *et al.*, 1991). Prominent members of the DBL family in *P. falciparum* are erythrocyte binding proteins EBA-140, EBA-175 and EBA-181. EBA-175 and EBA-140 bind to the RBC surface proteins glycophorin A and C respectively. The interaction partner of EBA-181 however has yet to be determined. DBL mediated surface binding is neuraminidase sensitive suggesting a

sialic-acid dependent interaction. PfRh homologues were first identified in *Plasmodium yoelii* and *P. vivax* and are believed to be involved in determining host-cell specificity (Galinski *et al.*, 1992;Preiser *et al.*, 2002). It appears that merozoites possibly invade with alternating receptors as transcription of *P. yoelii* reticulocyte binding homologues can vary within one single schizont and different variants can be expressed in different stages of the life-cycle (Preiser *et al.*, 2002). *P. falciparum* reticulocyte binding homologues are PfRh1, which binds to glycosylated RBC receptors and PfRh 2a/b and PfRh4 (Rayner *et al.*, 2001;Duraisingh *et al.*, 2003;Stubbs *et al.*, 2005). A novel RBC surface ligand for PfRh4 has recently been identified. Complement receptor 1 (CR1) interacts with PfRh4 in a sialic-acid independent invasion pathway (Tham *et al.*, 2010). Additionally, PfRh2b plays a role in invasion though its interaction partner has yet to be identified. EBA family members are thought to be released from micronemes whereas PfRhs are discharged from the rhoptry bulb prior to the irreversible attachment between merozoite and RBC.

The formation of the tight junction *per se* is poorly understood. AMA1 is translocated to the merozoite surface before it is shed, while some rhoptry neck proteins (RONs) are inserted into the host cell membrane. In *T. gondii*, AMA1 and the RON2/4/5/8 complex interact in an electron dense contact point of the moving junction which appears to surround the invading merozoite in a ring like structure (Alexander *et al.*, 2005;Lebrun *et al.*, 2005;Besteiro *et al.*, 2009;Besteiro *et al.*, 2011). This interaction between AMA1 and RON2/4 has recently been confirmed in *P. falciparum* (Cao *et al.*, 2009a;Collins *et al.*, 2009).

1.3.1.2 Invasion and PV formation

Upon the formation of a surface complex between AMA1 and RON2/4/5/8 *T. gondii* releases proteins and lipids into the host cell in order to establish the PV (Hakansson *et al.*, 2001). Similar cascades with discharge from rhoptries and dense granules have been suggested for *Plasmodium* (Aikawa *et al.*, 1983). The merozoite is ready and set to now push its way into the host cell under the calcium–dependent function of an actin-myosin motor. The motor is located just below the merozoite plasma membrane and is believed to interact with proteins of the surface complex (Green *et al.*, 2008). Shedding of merozoite surface adhesins that interact with RBC surface receptors from the anterior to posterior end is thought to help internalisation of the merozoite (Blackman, 2000;Harris *et al.*, 2005;O'Donnell *et al.*, 2005;O'Donnell *et al.*, 2006;Baum *et al.*, 2008); see Figure 6E). MSP1 and AMA1 are both shed by PfSUB2, a serine protease. A second family of serine proteases, the rhomboids which cleave in intra-membrane domains, are also involved in shedding of surface proteins. Rhomboids are believed to process surface adhesins like EBA-175, at the intra-

membrane and juxtamembrane domain in order to release the merozoite from its attachment to the host cell surface (O'Donnell *et al.*, 2005;Baker *et al.*, 2006;O'Donnell *et al.*, 2006). Although some understanding of how Apicomplexa enter new host cells has been collected, many questions remain, including the contribution made by proteolytic cascades in this complex process.

While invading the red blood cell the parasite creates the PV, a sack like compartment wherein the parasites resides and multiplies throughout its life-cycle. The PV is separated from the RBC cytosol by a porous PVM. The origin and assembly of the PVM is not clear. It is thought to contain both parasite and RBC components and its formation coincides with the release of the contents from the parasite's secretory organelles, such as dense granules and club-shaped rhoptries (Baum *et al.*, 2008) at the apical end into the host cytosol (Lingelbach *et al.*, 1998;Preiser *et al.*, 2000). While residing within the PV passive bidirectional transport across the PVM takes place to supply the growing parasite with nutrients (Nyalwidhe *et al.*, 2002). The exact purpose of the PV is yet to be established, although it has been shown that proteases and chaperones are the most abundant proteins in the PV lumen suggesting an involvement in protein sorting and nutrient uptake (Nyalwidhe *et al.*, 2006). The PV is a compartment of special interest particularly throughout the later parasite life stages. It plays a key part in the release of parasites into the blood stream as it is believed that proteases required for that step are discharged into the PV lumen just prior to egress (see above). Most Apicomplexan parasites establish a PV within the host cell**;** however, the timing of its persistence varies. *Toxoplasma* and *Plasmodium* reside within the vacuole throughout their entire intracellular stages whereas other Apicomplexa, for example *Theileria* and *Babesia* escape their PV shortly after invasion (Potgieter *et al.*, 1977;Shaw, 1997).

1.3.2 Getting out: Egress

At the end of asexual division the newly established daughter cells are required to escape from the host cell in order to continue the parasite life-cycle. This step is called egress (see Figure 7). The merozoites appear to become viable just prior to egress (Gilson *et al.*, 2009). Egress is essential in the asexual and sexual stages of the parasite's life-cycle. *P. falciparum* has to escape from hepatocytes, RBCs in the human host, RBCs after the uptake from the mosquito vector and the oocyst in the mosquito's midgut. The exact mechanism of egress is poorly understood and the proteins possibly involved have been under thorough investigation for two decades. An important role of proteolytic activity in disintegration of both membranes has been proposed (Deguercy *et al.*, 1990). Involvement of proteases was first shown by treating *in vitro* parasite cultures with a combination of several protease inhibitors. Parasite cultures treated with

the cysteine/serine protease inhibitor leupeptin, the broad spectrum protease inhibitor chymostatin, an aspartic protease inhibitor pepstatin and the cysteine/serine protease inhibitor antipain failed to rupture the host cell. Instead highly segmented and trapped schizonts could be observed (Lyon *et al.*, 1986).

To exit the infected host cell *P. falciparum* blood stage merozoites are required to breakdown two membranes, the PVM and the erythrocyte plasma membrane (EPM). Breaching the PVM might appear as the easier task as only one bi-layer has to be overcome (structural complexes have yet to be identified within the PV lumen). Overcoming the EPM might be considered more challenging with an intact cytoskeleton still in place. Early insights into egress were gained with the ambitious approach of recording the moment of egress of a duck malaria parasite (*Plasmodium iophurae*) with real-time live light microscopy (Trager, 1956), followed by recordings of *P. knowlesi* release from RBCs (Dvorak *et al.*, 1975). In the latter study an increase in intra-erythrocytic pressure was observed indicating that osmotic swelling might instigate egress in a mechanical way. Recordings of *P. falciparum* egress only emerged in 1999 (Winograd *et al.*, 1999;Wickham *et al.*, 2003;Glushakova *et al.*, 2005).

1.3.2.1 Current models of asexual blood-stage merozoite egress

Currently there are four major models describing egress of *Plasmodium* parasites (see Figure 8). Winograd *et al*. reported in 1999 (Winograd *et al.*, 1999) a non-explosive event which leads to parasite egress via localised fusion of the PVM and EPM (see Figure 8A). Merozoites were then slowly discharged into culture supernatant. This model is based on the observation that a membranous structure or RBC ghosts persists for a certain time period after the merozoites are released into the blood circulation. Densitometric measurement suggests that the RBC cytosol only slowly diffuses into the culture medium several seconds after the merozoites and residual body have left the host cell. A duct-like structure can be observed on the surface of the RBCs from which one or two merozoites escape at one time (Clavijo *et al.*, 1998).

A second model illustrates the hypothesis that the EPM undergoes degradation before PVM rupture (see Figure 8B). In a protease inhibitor study the membrane surrounding released merozoites was derived from the PVM (Salmon *et al.*, 2001). The resulting structures were referred to as so-called PVM-enclosed merozoite structures (PEMS). Parasites were cultured in the presence of N-[N-(L-3-transcarboxyirane-2-carbonyl)-L-Leucyl]-agmatine (E64) a covalent cysteine protease inhibitor, followed by colocalisation of merozoite PVM proteins and EPM proteins. According to the authors, this together with data from electron micrography of inhibited parasite cultures supports the hypothesis for an “outside-in” release of merozoites from erythrocytes.

A more recent study supports the idea that the PVM might remain intact around clusters of merozoites after the disintegration of the EPM (Soni *et al.*, 2005). The latter study took previous hypotheses further. Using IFAs and 3D reconstruction with antibodies specific to PVM or EPM proteins, the authors proposed that the PVM surrounds merozoites during late mitotic division hence inducing the formation of PEMS. At egress PEMS escape the RBC and only then the PVM ruptures under the control of cysteine proteases releasing invasive merozoites outside of the RBC. Notably both studies observe PEMS in untreated control *P. falciparum* cultures suggesting that they may represent egress intermediates rather than final egress structures, which contain merozoites and PVM residues. Additionally the reliability of information gained from IFAs can be questioned as the PVM was implicated to be of parasite and RBC origin. Therefore antibodies raised against RBC markers might not exclusively bind to the EPM but also recognise the PVM.

A third model suggests the breakdown of the PVM before disintegration of the EPM and is therefore a two-step egress also called the “inside-out” model (see Figure 8C). Wickham *et al.* investigated the mechanisms of egress by fluorescently tagging proteins which were exported either to the PV or the RBC cytosol of intra-erythrocytic *P. falciparum*. By differential temporal analysis with the help of protease inhibitors evidence was provided for a primary disintegration of the PVM before the EPM. In egress-inhibited parasites, both tagged proteins colocalise in the erythrocyte cytosol suggesting that PVM breakdown precedes EPM rupture. Additionally the authors proposed a role of distinct proteases being linked to distinct functions in egress as differential inhibition led to disintegration of either the PVM or EPM (Wickham *et al.*, 2003).

In accordance with the previous model Glushakova *et al*. described an explosive egress event following a cascade of parasite-induced modifications to the iRBC (see Figure 8D). All membranes were labelled non-specifically and in addition the EPM was biotinylated and detected with streptavidin-quantum dots and the relative localisation determined by microscopy. This approach is advantageous as it does not involve specific antibodies with a bias for altered binding properties between antibody and antigen as some intra-membrane proteins might undergo structural changes during egress. Next they followed egress by differential interference contrast (DIC) real-time microscopy. They showed that schizonts adopt a so-called “flower” formation immediately before rupture of the RBC. This stage is marked by the spatial distribution of the merozoites where they occupy the entire RBC cytosol. The formation of free PEMS could not be observed at any point. Glushakova *et al.* additionally attempted to investigate the membrane-fusion hypothesis (see above). By incubating iRBCs with positive-curvature inner leaflet amphiphiles an increase in parasitaemia should be

observed (if membrane breakdown during egress is facilitated by amphiphiles). Their results rule out the possibility of PVM and EPM fusion during egress. They suggest that modifications of the RBC cytoskeleton (possibly by proteases) and the increase in intracellular volume result in EPM rupture. The authors suggested that the breakdown of the PVM occurs immediately prior to or simultaneously with the breakdown of the EPM (Glushakova *et al.*, 2005). Recently, the same group very intriguingly showed the rupture of the PVM preceding the disintegration of the EPM, followed by an explosive release of merozoites (shown using confocal microscopy of living cells; Glushakova *et al.*, 2008;Glushakova *et al.*, 2010). In addition Glushakova *et al.* showed that the highly selective cysteine protease inhibitor E64 has an irreversible effect on the release of merozoites by affecting rupture of the EPM and not the PVM.

An additional approach to understand the order of egress was undertaken by analysing the sequential movement of the EPM during merozoite release. A succession of rapid movements by the EPM was described, beginning with curling of the EPM towards the outside and followed by buckling to retract the EPM even further. This eventually leads to eversion, the forceful release of merozoites and finally vesiculation of the remains of the EPM (Abkarian *et al.*, 2011;Lew, 2011). Other theories about egress have been suggested including the role of pore-forming proteins. The recent identification of the function of perforin-like proteins, or perforins, in *T. gondii* egress has led to the suggestion that pore formation is required for egress (Kafsack *et al.*, 2008;Kafsack *et al.*, 2010). Perforins contain a membrane attack complex (MAC) domain, which inserts into membranes forming pores, and are secreted by many cells for example natural killer cells and other immune cells (Horta, 1997). Perforin-like proteins were recently identified in *Plasmodium* (Kafsack *et al.*, 2010). The first indication of the role of perforins in *Plasmodium spp.* was given recently. Two perforins of *P. berghei* were shown to be indispensable in the parasite sexual stages as ookinetes with a disrupted membrane attack protein lost their infectivity in the mosquito (Andrea Ecker, unpublished; Ecker *et al.*, 2008)). The authors suggest that the ookinete lacks the ability to break down the midgut epithelium in order to complete the life-cycle (Kadota *et al.*, 2004). Perforins are considered to be actively involved in membrane degradation prior to parasite egress. Although present evidence of the temporal succession of PVM and EPM breakdown leaves room for speculation the necessity of protease participation is clear.

1.3.2.2 The role of proteases and calcium signaling in egress

The first insights into the function of proteases into egress were gained by approaches from two different perspectives; 1. the analysis of parasite cultures *in vitro* after the addition of specific inhibitors; 2. by identifying the activity and substrates of

specific proteases. A model for *P. falciparum* egress involving a cascade of protease activation has now been established with a cysteine protease called dipeptidyl peptidase 3 (DPAP3) and a serine protease called subtilisin like protease 1 (PfSUB1) as key players (Yeoh *et al.*, 2007;Arastu-Kapur *et al.*, 2008). PfSUB1 is released from dense granule-like organelles, termed exonemes (see Figure 5), into the PV lumen prior to egress where it processes many of its substrates (Janse *et al.*, 2007;Yeoh *et al.*, 2007;Silmon de Monerri *et al.*, 2011). The earlier mentioned MSP1 undergoes extensive primary processing by PfSUB1 prior to egress and invasion. In addition serine repeat antigen (SERA) family member SERA5, one of the most abundant proteins in the PV lumen, is processed by PfSUB1 just moments before egress at several identified processing sites (Yeoh *et al.*, 2007). Other substrates for PfSUB1 have recently been identified including proteins that are expressed late in the parasite life-cycle which suggests a role in egress and/or invasion (Silmon de Monerri *et al.*, 2011).

Previously the role of the PV was thought to be restricted primarily to the parasite's nutritional supply. However, recently its role in egress has been highlighted. The still intact or semi-intact PVM possibly ensures optimal temporal-spatial distribution of PV proteases and their substrates. This hints at the PV and PVM being involved in actively coordinating egress. Therefore putative proteases expressed late in the parasite life-cycle and secreted into the PV appear attractive targets for further investigation. In addition, certain parasite-derived proteases have been shown to be capable of processing RBC cytoskeleton components *in vitro* suggesting once again a role for proteases in egress (Deguercy *et al.*, 1990). However, it can be argued that the role of the PV evolved individually in different apicomplexan parasites as some egress without the presence of a PV (*Theileria* and *Babesia spp.*).

The exact mechanisms and the all important activators of egress however are mainly unknown. It remains unclear when and how fully developed merozoites receive the signal to breach the surrounding membranes in the hunt for fresh RBCs. Calcium-dependent signal transduction has been considered to play a key role in regulating this transition in the parasite life-cycle. The importance of calcium-mediated signalling was first shown by culturing *T. gondii* infected macrophages in the presence of calcium ionophores (Endo *et al.*, 1982) where protein kinases play a key role in invasion and possibly also egress. Recently, a plant like calcium-dependant protein kinase (PfCDPK5) has been identified in *P. falciparum*. The kinase is expressed in merozoites and is essential for parasite egress and survival, suggesting for the first time a direct involvement of calcium signalling in parasite escape from the erythrocyte (Dvorin *et al.*, 2010).

1.4 Proteases: a brief introduction

Proteases can be defined as enzymes that hydrolyse peptide bonds. Those enzymes cleave peptide bonds for an immense variety of specific or unspecific bio-chemical reactions to degrade, recycle or assemble peptide chains (Hooper, 2002). Proteases, also called peptidases or proteinases, are essential to the correct function of most organisms. Proteases represent about 2% of the total protein numbers of all organisms whose genomes have been sequenced to date (MEROPS database: http://merops.sanger.ac.uk). The human genome alone for example encodes ~500 proteases.

The catalytic process is mediated by the nucleophilic attack of either an essential active site residue or an activated water molecule on the carbonyl carbon atom of the scissile peptide bond. Hydrolytic enzymes are known to have two different types of active sites: either a catalytic dyad or catalytic triad protease domain, with two or three amino acid residues directly involved in the catalytic activity respectively. Some amino acid residues located within reach of the catalytic site usually deliver additional stabilising support. Protease families are based on the homology of the protease domain and several families have been identified: these include serine proteases, threonine proteases, cysteine proteases, aspartic proteases, glutamic proteases and metallo proteases (Rawlings *et al.*, 1993;Rawlings *et al.*, 2010). Metallo proteases received their name from the fact that they need to incorporate zinc ions through the side chains of the catalytic residues, called a zinc motif (Tallant *et al.*, 2010). Serine and cysteine proteases form a covalent acyl-enzyme intermediate during the reaction with active site residues facilitating the nucleophilic attack. In an aspartic acid- and metallo protease- mediated reaction the water molecule attacks the substrate rather than an active site residue. Proteases can function as endopeptidases to catalyse the degradation of peptide bonds between internal amino acids or as exopeptidases. Two different types of exopeptidases can be described that attack the polypeptide chain on its carboxy- or amino-terminus, named carboxy- and aminopeptidases respectively.

Proteases are categorised into clans, families and even more detailed groups depending on identity or similarity of the active side residues. Clans only have structural or sequence similarities due to evolutionary links whereas families are defined based on homologies. Serine proteases for example span many families and some of those have virtually identical structures or sequences establishing one clan (Rawlings *et al.*, 1993). Despite the identical catalytic site residues, proteases from within one type, e.g. serine proteases, vary widely in their amino acid composition, fold and substrate specificity.

It is important to distinguish between varying protease substrate specificities. Some cleave their substrate with a very high specificity and accuracy, allowing only very specific amino acid residues within the catalytic trough. Others accept a variety of residues within the substrate binding site with lower specificity. The active site of proteases has two functions, to recognise and bind a specific substrate and to catalyse the proteolytic processing. Residues within the active site or “groove” are generally labelled according to the Schechter and Berger nomenclature (Schechter *et al.*, 1967). Subsites (amino acids around the cleavage site within the substrate) upstream of the scissile bond are labelled P4-P1. Downstream residues are labelled P1′-P4′, where proteolytic processing occurs between P1 and P1′.

The best studied examples of less stringent protease activities with essential roles in the protein turnover of the cell are two mammalian cytoplasmic proteolytic pathways involving the proteasome and the lysosome (Clague *et al.*, 2010;Ciechanover, 2011). The lysosome is defined as a vacuolar structure with an acid pH containing degrading proteases, lipases and nucleases. It has been described as a non-specific degradative system (Whitford, 2005) and is mainly involved in the turnover of plasma proteins such as membrane channels. Aspartic, cysteine or serine proteases are the major components of the lysosome. Some of the serine proteases are members of the chymotrypsin family (Rawlings *et al.*, 1993). Chymotrypsin and trypsin are promiscuous serine proteases with major roles in many organisms ranging from viruses to vertebrates (Rawlings *et al.*, 1994;Barrett *et al.*, 1995). In vertebrates for example chymotrypsin is secreted into the digestive tract to process peptides so that they can be absorbed from the gut endothelium.

A second pathway with importance in the cellular protein turnover is the proteasome. In contrast to the lysosome, the proteasome is comprised of ATP-dependent endopeptidases with multicatalytic domains forming a barrel shaped complex of 2.1 mega Dalton (mDa). The proteasome degrades 90% of all cellular proteins (Whitford, 2005). The eukaryotic 26S is the best known proteasomal structure composed of two 19S cap regulatory units and one 20S core barrel shaped structure (Peters *et al.*, 1994). The 20S proteasome is thought to act in a manner similar to threonine-proteases and in most cases proteolytically processes proteins that have been ubiquitinated (Voges *et al.*, 1999;Rape *et al.*, 2002). The above mentioned activity patterns highlight the complexity involved in proteolytic processing in a high number of different organisms.

1.5 *Plasmodium* proteases in a wider context

Proteases have been shown to play major roles in the intracellular development of apicomplexan parasites. Many proteolytic enzymes have been identified and

characterised in *P. falciparum*, *T. gondii* and other Apicomplexa. It has long been thought that *P. falciparum* requires protease activity in order to egress, invade or proliferate. A total of 92 proteases have been identified in the genome of *P. falciparum*, however the function of the majority of these is unknown (Wu *et al.*, 2003). Vanguard studies using protease inhibitors proved to be milestones in understanding parasitic protozoan biology. They showed that broad spectrum protease inhibitors arrest both *P. falciparum* schizont egress (Lyon *et al.*, 1986) and the invasion of rhesus monkey erythrocytes by *P. knowlesi* merozoites (Hadley *et al.*, 1983). These approaches were the first opportunity to gain insights into protease regulated mechanisms of protozoan parasites. Shortly after, direct implications for the involvement of proteases in specific erythrocytic stages were made (Rosenthal *et al.*, 1987). Several mechanistic classes of *Plasmodium* proteases have been identified. Since then many approaches with various serine, cysteine and other protease inhibitors showed the involvement of proteolytic activity in egress and invasion of *P. falciparum* merozoites and in other extra-erythrocytic stages (Blackman, 2008). Their exact function however is poorly understood. With many proteins undergoing essential proteolytic processing throughout distinct life-cycle stages proteases are now one of the main targets for innovative antimalarial research.

1.5.1 Metallo proteases

P. falciparum metallo proteases are very poorly understood. In order to accomplish substrate recognition and catalytic activity metallo proteases require the binding of metal ion cofactors such as zinc to a specific binding motif. The imidazole side chains of two His residues within the motif confer the interaction with Zn^{++}. Once bound to the imidazole ligands, zinc promotes the nucleophilic attack on the peptide bond. Metallo proteases are believed to be involved in parasite invasion as the treatment with metal chelators leads to egress but inhibits invasion (Kitjaroentham *et al.*, 2006). In 2003 Wu *et al*. identified several previously unknown *P. falciparum* metallo proteases, preparing the ground for future work (Wu *et al.*, 2003). Only one *P. falciparum* metallo protease has been described and its action investigated. Falcilysin, a member of the M16 metallo protease family is expressed in erythrocytic stages of *P. falciparum* parasites. It was first described to be involved in degradation of small haemoglobin derived peptides within the acidic food vacuole (Eggleson *et al.*, 1999). However, it was noted that falcilysin alters substrate specificity depending of the surrounding pH. Additionally the metallo protease was not only distributed within the acidic food vacuole but also outside within other parasite organelles, suggesting a dual function (Murata *et al.*, 2003). Falcilysin has been hypothesised to also be involved in degrading apicoplast transit peptides (Ponpuak *et al.*, 2007).

1.5.2 Aspartic Proteases

Aspartic proteases exist in an extensive host range from viruses to plants, and fungi to mammals. Humans are thought to have used the aspartic protease chymosin in the form of rennet for thousands of years in cheese production. They were the first to be described and are well understood in their function, e.g. porcine pepsin (Szecsi, 1992). An acid/base catalytic activity is conferred by two aspartic acid residues, one functioning as a proton acceptor from a water molecule-derived hydrogen. A water molecule is activated by general acid-base catalysis, and then acts as the nucleophile (Coombs *et al.*, 2001). It is important to mention that aspartic proteases do not form a covalent acyl-enzyme intermediate. Not only are aspartic proteases well characterised in humans but they are also the best studied proteases within the Apicomplexa. The function of at least five aspartic proteases has been identified and characterised in *P. falciparum*. The major group of aspartic proteases in *P. falciparum* is the plasmepsins. Ten of these can be found in the genome of the parasite and appear to have orthologues in most *Plasmodium* species (Gardner *et al.*, 2002;Shea *et al.*, 2007). Plasmepsin I and plasmepsin II were the first members of the aspartic protease family to undergo thorough functional analysis. Together with two plasmepsin I and II homologues, plasmepsin IV and hispo-aspartic protease (HAP), they are located within the food vacuole (FV) and synthesised as type II integral membrane proteins with the active site placed within the FV. Once inserted into the membrane, falcipains, members of the cysteine protease family (see 1.5.5.2), activate plasmepsins by cleaving off the inhibitory propeptide and releasing a now soluble and active protease into the FV lumen (Drew *et al.*, 2008). This processing step is redundant as in transgenic parasites which lack falcipains, plasmepsins undergo autocatalytic processing (Drew *et al.*, 2008). HAP, which is also named plasmepsin III, is an interesting aspartic acid protease as it carries an active-site His residue substitution (Berry *et al.*, 1999;Humphreys *et al.*, 1999). However, this replacement does not appear to cause any impairment in catalytic activity (Banerjee *et al.*, 2002).

It is essential for the parasite to digest RBC haemoglobin within the FV in order to gain amino acids that in return can be incorporated into protein synthesis (Sherman *et al.*, 2004). The FV-located plasmepsins are responsible for haemoglobin degradation in what appears to be a regulated cascade involving first aspartic proteases followed by cysteine proteases, then metallo proteases, and finally aminopeptidases (Gluzman *et al.*, 1994;Banerjee *et al.*, 2002;Klemba *et al.*, 2002). HAP and plasmepsin IV are believed to act downstream of plasmepsin I/II (Banerjee *et al.*, 2002). Although not individually essential, the genetic disruption of each of plasmepsin I-IV individually in the parasite led to reduced growth rates of the transgenic lines. Simultaneous deletion of all four of these plasmepsins leads to deformation and malfunction of the FV (Bonilla

et al., 2007a;Bonilla *et al.*, 2007b). Their roles are partially overlapping which suggests functional redundancy and might also indicate why *P. falciparum* maintained four copies in the genome (Omara-Opyene *et al.*, 2004;Liu *et al.*, 2005;Bonilla *et al.*, 2007a;Bonilla *et al.*, 2007b). Interestingly genetic disruption of plasmepsin IV in the rodent parasite, *P. berghei*, confers reduced virulence and protective immunity against experimental malaria which is possibly due to lower fitness (Spaccapelo *et al.*, 2010). A similar role for *P. falciparum* plasmepsins cannot be excluded. Additionally, in *T. gondii* aspartic protease 1 (a possible homologue of plasmepsin I) is suggested to have a role in cell division, indication further and yet unexplored roles for the *P. falciparum* plasmepsins (Shea *et al.*, 2007). In a first instance it was unclear whether plasmepsins could serve as new antimalarial drug targets because of their similarity to cathepsin D and E, two major and very important human aspartic proteases. However, thermodynamic studies showed that in fact both proteases significantly differ in their activity, which opens doors for the design of specific plasmepsin inhibitors (Xie *et al.*, 1997). But the previous functional redundancy has to be considered in the attempt to design and screen for novel anti-plasmepsin drugs.

The role of the other six aspartic proteases remains less well characterised. Plasmepsin V resides within the endoplasmatic reticulum (ER; Klemba *et al.*, 2005). It has only very recently been characterised as an essential protease, which proteolytically processes cargo proteins downstream of a certain export signal called the PEXEL motif (Boddey *et al.*, 2009;Russo *et al.*, 2009a;Haase *et al.*, 2010). Plasmepsin VI/VII/VIII are not expressed in asexual parasite stages and are yet to be characterised (Florens *et al.*, 2002;Le Roch *et al.*, 2004). Plasmepsin IX and X are both expressed late in asexual stages indicating a putative role in parasite egress or invasion. Plasmepsin X was found in a yeast two-hybrid system to interact with *P. falciparum* asparagine and aspartate rich protein 2 (PfAARP2) (LaCount *et al.*, 2005). PfAARP2 is believed to localise to the RBC cytosol (Barale *et al.*, 1997). This might be considered particularly interesting as plasmepsin II was shown to process the erythrocyte skeleton proteins actin, band 4.1 and spectrin *in vitro* (Le Bonniec *et al.*, 1999), suggesting that some plasmepsin substrates could be found in the RBC.

1.5.3 Serine proteases: Subtilisins

Serine proteases are endopeptidases that constitute the largest and most diverse group of proteolytically active enzymes. First insights into the understanding of the active site were gained by the analysis of chymotrypsin in the 1960s (Bender *et al.*, 1973). Chymotrypsin was also the first protease for which a three-dimensional structure was solved (Matthews *et al.*, 1967;Matthews *et al.*, 1977). Serine proteases are found in eukaryotes, prokaryotes, archaea and viruses where they are involved in processes

as diverse as digestion and immune responses (Polgar, 2005). According to their degree of homology serine proteases are classified into four major clans. They are then further categorised into families in agreement with their substrate specificity and activity spectrum. The two best understood and characterised clans are the chymotrypsin-like and subtilisin or subtilisin-like proteases. The chymotrypsin-like family includes promiscuous proteases such chymotrypsin, trypsin and elastase. In humans all three proteases are produced and secreted within and from the pancreas and play major roles in the digestive system. Although they all appear similar in structure the substrate specificity varies considerably. Chymotrypsin prefers to proteolytically cleave after a rather large and hydrophobic residue. Trypsin processes most successfully and specifically after a positively charged amino acid. Elastase on the other hand requires a small neutral residue in order to fully cleave a peptide bond. These three examples highlight the diversity within the serine proteases. The second, and, from the *P. falciparum* view, the most important clan are the subtilisins or subtilisin-like proteases. Subtilisins derived their name from the organism in which they were first discovered and from which they were first isolated, *Bacillus subtilis*. Under conditions of nutrient deprivation these bacteria express subtilisins in high amounts (Pierce *et al.*, 1992). Interestingly subtilisins and the chymotrypsin-like proteases share the same catalytic mechanisms, even though sequence and structural identity between both is exceedingly low. Subtilisins are well-characterised and commercially applied in somewhat unusual methods such as additives in many cosmetics and cleaning detergents (Aaslyng *et al.*, 1991). There are over 200 known subtilisins and based on their sequence identity they are divided into six families (Siezen *et al.*, 1997). All subtilisins require Ca^{2+} for stability and activity. Activity is conferred by a classical catalytic triad which is comprised of a His-, a Ser- and an Asp residue. The active site Ser residue undergoes activation by the imidazole ring of the His residue. The imidazole accepts a hydrogen (proton) from the serine in order to enable the hydroxyl group to begin the nucleophilic attack on the carbonyl carbon of the scissile peptide bond to form a tetrahedral intermediate. The imidazole ring however can only accept a hydrogen proton if the Asp residue functions as electron donor to begin with. The incorporation of a water molecule then leads to the completion of the reaction and release of the substrate. Additional residues are required in order to stabilise the reaction.

The essentiality of the catalytic Ser of serine proteases has been shown by site directed mutagenesis that only alters the Ser to an Ala residue. Mutations to the Ser residue inhibit the proteolytic processing of substrates (Carter *et al.*, 1988). In the same study it was shown that the mutation has no impact on the actual protease fold and only very little impact on the substrate affinity. Therefore replacing the catalytically

essential residue with a “silent” Ala residue has since been used to study protease activity for many peptidases. Subtilisins are generally expressed as inactive protease precursors, called zymogens or proproteins, which carry an inhibitory prodomain. Upon autocatalytic processing subtilisins are separated from the prodomain and the active protease domain is released. One important and well characterised example of autocatalytic activation is furin a subtilisin-like serine protease which undergoes ordered, pH dependent and compartment specific auto-processing which leads to its activation (Anderson *et al.*, 1997;Anderson *et al.*, 2002). Folding of furin occurs in the ER where it also undergoes the first initial auto-cleavage. This step is not sufficient to separate the prodomain from the active protease and together they are trafficked to the trans-Golgi network. Now entering an acidic compartment a second autocatalytic step takes place to free the now fully active protease from its inhibitory prodomain (Anderson *et al.*, 1997). Prodomains are considered to function in many ways. They are extremely specific inhibitors and intramolecular chaperones that ensure the correct fold and accurate trafficking of the inactive protease precursor (Anderson *et al.*, 2002). The hypothesis of protease folding, directed inhibition and trafficking to the place of activity is now widely accepted as a general model for the activation of subtilisins and proteases of other clans (Yabuta *et al.*, 2003;Subbian *et al.*, 2005).

1.5.3.1 *P. falciparum* subtilisins

As mentioned above, broad spectrum protease inhibitor studies led to the conclusion that serine proteases are essential for parasite egress from and subsequent invasion of RBCs (Dluzewski *et al.*, 1986;Blackman *et al.*, 1993;Blackman, 2000). The *P. falciparum* genome encodes three subtilisin-like proteases or subtilisins, termed PfSUB1, 2 and 3 (Withers-Martinez *et al.*, 2004). Therefore these proteases appeared as prospective candidates for fulfilling important functions within the above mentioned life-cycle stages. PfSUB1 was the first of these subtilisin-like proteases to be described (Blackman *et al.*, 1998). The first characterisation of PfSUB2 followed shortly after, by approaches from two independent laboratories (Barale *et al.*, 1999;Hackett *et al.*, 1999). Both PfSUB1 and PfSUB2 are active proteases, and PfSUB3 displays weak autocatalytic processing in recombinantly expressed material (Michael Shea, unpublished). All three subtilisin-like proteases only exist as a single copy within the genome. Expression of PfSUB1 and PfSUB2 peaks at very late schizont stages making those proteases even more intriguing as to whether they play important roles in egress and invasion (Blackman *et al.*, 1998). PfSUB1 and PfSUB2 are refractory to deletion in the parasite indicating an essential function. Despite sharing a conserved subtilisin-like protease domain the protein composition and fold of PfSUB1 and PfSUB2 varies remarkably. PfSUB1 is a soluble protein with a rather small size of 688 amino acid

residues. PfSUB2 however is much larger comprising twice as many amino acids as PfSUB1 and contains at least one membrane spanning domain downstream of the catalytic groove. It is considered a type I integral membrane protein.

Several subtilisin-like proteases have been identified in other apicomplexan species, e.g. *Babesia divergens*, *T. gondii* and *Neospora caninum*. All *Plasmodium* species have one copy of SUB1 and SUB2. Out of all these orthologues only SUB2 has been studied in *P. berghei*, where it is expressed in merozoites (Han *et al.*, 2000;Uzureau *et al.*, 2004).

1.5.3.2 PfSUB1

PfSUB1 is expressed as a proteolytically inactive zymogen. It requires auto-processing in order to become active and hydrolyse its substrates. PfSUB1 activity was first shown by recombinant expression of the protease followed by autocatalytic processing (Sajid *et al.*, 2000). Several approaches with fluorogenically labelled peptides based on the autocatalytic cleavage site then revealed further insight into PfSUB1 substrate specificity and protease activity (Sajid *et al.*, 2000;Blackman *et al.*, 2002;Withers-Martinez *et al.*, 2002;Jean *et al.*, 2003). Based on broad spectrum protease inhibitor studies, PfSUB1 was first postulated to be responsible for shedding the merozoite surface proteins MSP1 and AMA1 prior to and during invasion. However, specific protease inhibitors of PfSUB1 do not interfere with the shedding of these proteins but do inhibit invasion (Blackman, 2004). These results excluded the theory that PfSUB1 might be displaying the previously predicted sheddase function. PfSUB1 biological activity and function has since been investigated and understood in much more detail. It exhibits not only a role in invasion but also in egress. In order to gain insights into PfSUB1's biological role, its location and native substrate had to be identified. IFA revealed that PfSUB1 does not colocalise with any markers of known *P. falciparum* compartments such as RhopH2 (rhoptry protein), AMA1 (microneme protein) or RESA (dense granule protein). Immunoelectron microscopy (EM) suggested that PfSUB1 localises to a dense granule like organelle which is however distinct from dense granules, as proteins such as RESA do not colocalise with PfSUB1 even in EM (Aikawa *et al.*, 1990;Harris *et al.*, 2005;Yeoh *et al.*, 2007). These novel organelles have now been named exonemes (Janse *et al.*, 2007). Additionally Yeoh *et al.* discovered that selective PfSUB1 inhibitors interfere with *P. falciparum* egress and block the proteolytic processing of the highly abundant putative papain-like protease SERA5. SERA5 was identified as the first physiological substrate of PfSUB1 and the mapping of the proteolytic processing sites established an improved picture of PfSUB1 substrate specificity and function. SERA5 is a well characterised protein which is located in the PV (see 1.6.2). Therefore PfSUB1 is believed to possibly induce a proteolytic cascade

facilitating schizont rupture and the release of merozoites from the host RBC. Exoneme discharge precedes egress with PfSUB1 translocating to the PV where it now meets its physiological substrates such as SERA5 (Yeoh *et al.*, 2007).

PfSUB1 possibly also exhibits a role in merozoite invasion by processing merozoite surface proteins in late schizogony, so as to prime them for invasion. Indeed, at lower concentrations, the specific inhibitor of PfSUB1 does not block egress but still inhibits invasion (Yeoh *et al.*, 2007;Koussis *et al.*, 2009). The complex between MSP1, MSP6 and MSP7 uniformly coats the surface of merozoites and is suggested to be involved in receptor mediated interactions that precede merozoite invasion (Goel *et al.*, 2003;Li *et al.*, 2004;Kauth *et al.*, 2006). This MSP complex is known to undergo several proteolytic cleavages in the parasite prior to invasion. Recombinant PfSUB1 cleaves parasite-derived MSP1/MSP6/MSP7 *in vitro* at the same sites that they are processed in the parasite (Kadekoppala *et al.*, 2008;Kadekoppala *et al.*, 2010). Additionally substrate peptides based on these sites are also correctly processed by recombinant PfSUB1. To confirm these results a specific and highly efficient inhibitor of PfSUB1 was required (the chemical inhibitor described above is from a relatively impure extract and shows activity only in the micromolar range). A characteristic of subtilisin-like proteases is that the prodomain generally forms a specific and high-affinity inhibitor of its cognate enzyme. The recombinant PfSUB1 prodomain indeed proved to be a specific nanomolar inhibitor of recombinant PfSUB1 (Jean *et al.*, 2003) and allowed confirmation of the cleavage of MSP1/6/7 by recombinant PfSUB1 *in vitro* (Koussis *et al.*, 2009). This processing step however displays priming properties rather than shedding as the complex remains associated with the surface of the merozoite. A recent global proteomic analysis identified additional PfSUB1 substrates and a significantly improved model for PfSUB1 substrate specificity was compiled (Silmon de Monerri *et al.*, 2011). PfSUB1 prefers a hydrophobic residue in the P4 position and a small residue in the P2 position (Gly, Ala); P1 is preferably polar and all P′positions should have an acidic tendency (see Figure 9). Alignments of SERA5 with all the other members of the SERA family indicated that other members, e.g. SERA6, putatively undergo processing by PfSUB1 (see 1.7). Despite extensive efforts a crystal structure of PfSUB1 is unfortunately not yet available.

1.5.3.3 PfSUB2

According to homology modelling of the active site groove of PfSUB2, this protease was proposed as a candidate for being responsible for the secondary processing and shedding of MSP1 during merozoite invasion (Barale *et al.*, 1999). First insights into PfSUB2 activity and substrate specificity however were not gained before 2005 when Harris *et al*. were able to identify the protease as the “sheddase” active

during merozoite invasion. In a very sophisticated approach PfSUB2 recombinant prodomain was incubated with parasite cultures in order to bind to and inhibit parasite derived mature PfSUB2. This assay indeed revealed PfSUB2-mediated shedding of the essential surface proteins MSP1 and AMA1. PfSUB2 was found to be secreted from micronemes during egress onto the anterior end of the merozoite surface. Upon invasion PfSUB2 translocates to the posterior pole of the merozoite. During this movement PfSUB2 “sheds” the ectodomains of MSP1 and AMA1 to disengage adhesin-complexes (Harris *et al.*, 2005) and allows the merozoite to become engulfed by the RBC. Very recent unpublished data suggest that AMA1 shedding occurs in a sequence unspecific manner. The recognition motif appears to comprise the distance from the transmembrane domain rather an actual amino acid sequence (Anna Olivieri, Christine Collins and Michael Blackman, unpublished). A third physiological PfSUB2 substrate was recently identified. Plasmodium thrombospondin related apical merozoite protein (PTRAMP) is located within the micronemes and relocated to the apical pole of the merozoite to then redistribute itself onto the entire surface upon invasion. Its role is unclear but involvement in invasion has been suggested. After its translocation PTRAMP undergoes shedding by PfSUB2 during invasion (Green *et al.*, 2006). Despite extensive efforts recombinant PfSUB2 is not yet available.

Early protease studies have shown that two different classes of serine proteases are required for invasion which leads on to the *P. falciparum* rhomboid proteases (Hadley *et al.*, 1983).

1.5.4 Serine proteases: Rhomboids

An additional rather unusual group of serine proteases, the rhomboids has been recently identified in the fruit fly *Drosophila melanogaster*. Their uniqueness is conferred by the fact that rhomboids contain six to seven intramembrane domains and are capable of proteolytically cleaving substrates within a phospholipid bilayer or in juxtamembrane domains (Urban *et al.*, 2001;Freeman, 2008;Freeman, 2009). Additionally unlike the other serine-type proteases, rhomboids use a catalytic dyad and are very conserved in eukaryotes and prokaryotes (Urban *et al.*, 2002a;Urban *et al.*, 2002b).

1.5.4.1 *P. falciparum* rhomboids

Apicomplexan rhomboids (ROMs) are a well studied group of proteases. Studies in *T. gondii* give the best understood insight into the function of this family (Kim, 2004;O'Donnell *et al.*, 2005). A role of rhomboids in invasion by protozoan parasites was first suggested in 2003 as a group of putative substrates was identified on the surface of *T. gondii* (Urban *et al.*, 2003). The first experimental approach that

deepened the idea for a role for rhomboids throughout these steps was conducted by applying proteomic analysis of cleavage events prior to and during invasion (Zhou *et al.*, 2004). Surface adhesin proteins, called TgMICs, which are released from *T. gondii* micronemes onto the parasite surface moments before parasite egress (Carruthers *et al.*, 1999a), were shown to be processed by rhomboids. TgMIC2, TgMIC6 and TgMIC12 are proteolytically cleaved within their transmembrane domain by a protease called *T. gondii* microneme protein protease 1 (TgMPP1) (Carruthers *et al.*, 2000;Opitz *et al.*, 2002). TgMPP1 is conserved throughout Apicomplexa and encoded by a rhomboid gene. An uncleavable TgMIC2 mutant indicated that TgMPP1 processing is essential for parasite invasion (Brossier *et al.*, 2003). It has now been suggested that TgROM5 might be the as yet unidentified MPP1 (Brossier *et al.*, 2005). Combined with the first evidence listed above, genome wide analysis revealed that rhomboid genes are found in all apicomplexan parasites. *P. falciparum* encodes eight rhomboids (PfROM1, 3, 4, 6, 7, 8, 9, 10) with four being orthologues of *T. gondii* rhomboids (Dowse *et al.*, 2005). PfROM1 and PfROM4 were shown to cleave a variety of substrates containing transmembrane domains and are possibly involved in all invasive stages of *P. falciparum* (Baker *et al.*, 2006). PfROM4 processing of EBA-175 is essential as parasites expressing a cleavage site mutant cannot survive (O'Donnell *et al.*, 2006). ROM1 is thought to be involved in the intracellular development of parasites. *T. gondii* parasites with conditionally knocked-down TgROM1 replicate significantly more slowly within the host cell (Brossier *et al.*, 2008). In addition mutants of the rodent parasite *P. berghei* with PbROM1 knocked-out are incapable of forming oocysts and display a decreased hepatocyte invasion rate (Srinivasan *et al.*, 2009). An extended role for rhomboids has recently been suggested as *T. gondii* AMA1 processing by TgROM4 induces a switch from invasive to replication mode (Santos *et al.*, 2011). Finally, it has to be mentioned that although broad spectrum protease inhibitors reduce rhomboid activity no specific inhibitors have been identified to date (Urban *et al.*, 2001). Rhomboids are clearly a very interesting, distinct and essential class of active serine proteases.

1.5.5 Cysteine proteases

Cysteine proteases perform an extensive variety of functions. They can be found in many organisms ranging from viruses to plants to humans. Their roles in plants involve proteolytic activity in growth, development and programmed cell-death, and therefore in many signalling pathways (Grudkowska *et al.*, 2004). In humans, cysteine proteases are involved in protein degradation (lysosomes or proteasomes), bone resorption, and parts of the immune response and are implicated in several diseases such as proliferation of metastatic cells in cancer (Nomura *et al.*, 2005). Many

parasitic diseases such as malaria are reliant on functional and regulated cysteine protease activity (McKerrow *et al.*, 1999;McKerrow *et al.*, 2006). The catalytic activity of cysteine proteases is similar to serine proteases with the Ser replaced by a Cys residue. All cysteine proteases are subdivided into clans. Similarity between different clans is very low although they are suggested to have developed from the same origin (MEROPS database: http://merops.sanger.ac.uk). Each clan is further sub-divided into families this time based on sequence and structural homologies. Probably the best known cysteine protease was first found as the active constituent in the latex of the fruit *Carica papaya* (papaya) and was therefore named papain. Papain formed the foundation of the protease clan CA. Clan CA, which includes all papain-like protease families, e.g. cathepsins (family C1) and calpains (family C2), and clan CD, including caspases (family C14) are the best studied cysteine protease clans (Barrett *et al.*, 2001;Rawlings *et al.*, 2010). Although they all share a conserved catalytically active cysteine, activity is conferred by catalytic triads or catalytic dyads respectively for CA and CD with a Cys and a His residue forming the core of the protease motif. As described earlier for serine and other proteases (see 1.5.3), cysteine type proteases are usually expressed as zymogens or proproteins for precise regulation of proteolytic activity. The conserved protease activity depends on the thiol group of the Cys residue undergoing deprotonation by the imidazole ring of the His residue. The nucleophilic attack on the carbonyl carbon of the peptide bond is in then performed by the deprotonated anionic sulphur, forming a thioester intermediate between the new carboxy-terminus and Cys thiol. The incorporation of a water molecule induces the release of the new carboxy-terminus of the substrate and the regeneration of all the active site residues to their original electron balance. All proteases within clan CA (papain-like proteases) possess a catalytically active triad with the involvement of an additional Gln residue. The papain-like protease domain therefore comprises as follows: Gln, Cys, His and Asn/Asp (MEROPS database: http://merops.sanger.ac.uk). Other proteases such as caspases only require a catalytic dyad for full activity.

1.5.5.1 The role of cysteine proteases in *P. falciparum*

Sequencing of an ever increasing amount of *Plasmodium spp.* genomes and development of techniques to manipulate the parasites genomic information are probably the most influential advances in order to gain information about malarial cysteine proteases (Crabb, 2002). The first and most promising insights were derived from the previously mentioned protease inhibitor studies. These studies examined the function and role of cysteine proteases as a group throughout the parasites life-cycle rather than identifying a specific function (Rosenthal *et al.*, 1987). For example the inhibition of different life-stages of *P. falciparum* cultures with cysteine protease

inhibitors such as leupeptin and E64 arrested parasite development. When analysed in greater detail it was observed that processes such as haemoglobin degradation were abolished (Shenai *et al.*, 2002). Since the first studies emerged *P. falciparum* cysteine proteases have been shown to be involved in egress, invasion and growth and replication (Rosenthal *et al.*, 1988;Greenbaum *et al.*, 2002;Sabnis *et al.*, 2003;Singh *et al.*, 2006;Wang *et al.*, 2011). Many more putative cysteine proteases, e.g. the SERAs (papain-like), calpains and cathepsins have since been identified in the *P. falciparum* genome with some currently undergoing thorough investigation (Wu *et al.*, 2003).

1.5.5.2 *P. falciparum* falcipains

Falcipains were first biochemically characterised as members of the papain clan (Gln-Cys-His-Asn active site), localised to the FV and named trophozoite proteinases as their cysteine protease activity was revealed in trophozoite stages (Rosenthal *et al.*, 1988). Protease inhibitor studies clearly indicated the importance of cysteine proteases in haemoglobin digestion as the inhibition of these proteases led to extensive swelling of the FV in parasite cultures (Bailly *et al.*, 1992). Soon after the first gene was characterised; the protease was named falcipain and it was shown to be a haemoglobinase (as are the plasmepsins; Rosenthal *et al.*, 1992). Since then three additional falcipains have been identified. There are four falcipains in total: falcipain 1, which is encoded on chromosome 1, falcipain 2, falcipain 2′ and falcipain 3 all of which are encoded on chromosome 14 (Rosenthal, 2004). Falcipain 2 and falcipain 2′ share 99% identity in catalytic regions and are expressed in earlier parasite asexual life stages (Singh *et al.*, 2006). Falcipain 3, despite the ability to degrade haemoglobin *in vitro*, undergoes upregulation in the middle of the parasites erythrocytic development (Sijwali *et al.*, 2001b). Experimental knockouts of falcipain 1 led to the conclusion that asexual RBC stages of *P. falciparum* do not require the protease to survive (Sijwali *et al.*, 2004a). Genetic disruption leading to functional knockouts of falcipain 2 however resulted in a swollen FV indicating the impairment of haemoglobin digestion (Sijwali *et al.*, 2004b). The knockout can be rescued by the expression of an additional, episomal copy of falcipain 2 (Armstrong *et al.*, 2007). Interestingly in the falcipain 2 knockout parasites the nearly identical falcipain 2′ does not undergo upregulation. This appears surprising as falcipain 2 knockout lines, despite displaying a swollen FV, develop normally (Sijwali *et al.*, 2004b). Additionally falcipain 2 also digests ankyrin and band 4.1 at neutral pH *in vitro* (Dua *et al.*, 2001;Hanspal *et al.*, 2002). In combination with the fact that a peptide that inhibits recombinant falcipain 2 also inhibits egress and that falcipain 2 is also localised to the PV (Dhawan *et al.*, 2003), the role of this cysteine protease appears to not be as well understood as previously thought and a function in egress cannot be excluded. Falcipain 2 also dispensable in the parasites

erythrocytic life stages (Blackman, 2008). Falcipain 3 on the other hand is refractory to genetic manipulation and is upregulated in falcipain 2 knockout lines. Its function, however, still remains unknown (Singh *et al.*, 2006). All falcipains are expressed as proproteins with falcipain 2 displaying an unusually long prosequence (Shenai *et al.*, 2000). Falcipains appear to be involved in proteolytic cascades leading to haemoglobin hydrolysis as they serve as maturases for plasmepsins within the FV lumen (Drew *et al.*, 2008). Despite falcipain 1, 2 and 2′ appearing to be dispensable for the parasite the essential nature of falcipain 3 makes this haemoglobinase an intriguing target for antimalarial development (Kerr *et al.*, 2009;Teixeira *et al.*, 2011).

1.5.5.3 *P. falciparum* calpains

Calpains (calcium ion-dependent papain-like cysteine protease) are Ca^{2+}-dependent multifunctional cysteine proteases which usually show regulated activity at a neutral pH. Calpains are ubiquitously expressed in birds and mammals and they are thought to be involved in cell cycle regulation and cytoskeletal remodelling. The *P. falciparum* calpain homologue is very poorly understood and is refractory to genetic disruption. This has led to the suggestion that this homologue displays an essential function for the parasite (Russo *et al.*, 2009b). *P. falciparum* cultures maintained in the presence of specific calpain inhibitors experience significantly reduced merozoite invasion rates (Olaya *et al.*, 1991). Of course an effect of the inhibitors on the RBC calpain, calpain-1, cannot be excluded. Intriguingly if RBCs undergo chemical- and immunodepletion of calpain-1, parasites are incapable of egress. However, after egress has been restored merozoites invade less efficiently (Chandramohanadas *et al.*, 2009).

1.6 SERA family

Two of the most abundant protease families in *P. falciparum* include the falcipains (papain-like cysteine proteases) and the aspartic proteases of the plasmepsin family. These were shown to mainly be involved in degradation of haemoglobin, but may also play a role in degradation of the RBC cytoskeleton (Rosenthal, 2002;Rosenthal, 2004). An additional group of putative *Plasmodium spp.* cysteine proteases is made up of the members of the SERA family. SERAs are composed of a papain-like central region with long N- and C-terminal extensions. The name of the SERA family is derived from a long stretch of Ser residues within the most highly-expressed (and first-identified) member of this group, SERA5 (Bzik *et al.*, 1988). The SERA multigene family is nearly unique to *Plasmodium*, as only *Theileria* expresses one orthologue.

In recent years efforts have been made to investigate the origin of the SERAs. All *Plasmodium* species encode SERAs in their genomes. They are usually found in gene clusters. However, numbers of *sera* genes range from two in *Plasmodium gallinaceum* to 12 in *P. vivax*. The SERA family has recently been divided into four groups, depending on their putative fold and sequence homology (Arisue *et al.*, 2011). Intriguingly within these groups there are two different types of SERAs. Some contain a Cys residue in the active site (groups I/II/III), like papain. Some however have the Cys substituted with a Ser residue (group IV). The evolutionary process causing the establishment of the unique SERAs and the substitution of a Ser in the active site remains unknown. As no SERA protease activity has convincingly been shown for any SERA family member it is difficult to speculate about the effect of the active site substitution. It is very interesting to observe that SERAs of the mammalian *Plasmodia* vary in numbers only in group IV, the serine type. All *Plasmodium spp.* have at least one member of group I to III, with the one exception of *P. gallinaceum* which encodes one member of group I and one ancestor of group II/III (Arisue *et al.*, 2007;Arisue *et al.*, 2011). The gene number variation occurred most strongly in human, ape and monkey groups possibly indicating that adaption to different hosts led to evolution of the SERAs (see Figure 10, Table 1). Additionally some members of group IV have lost the central-papain like domain. This might indicate that the serine-type SERAs do not require the papain-like domain to achieve their complete function.

1.6.1 The SERA family of *P. falciparum*

The SERA family was first discovered in the 1980s in an attempt to identify novel parasite antigens as vaccine candidates. In the very first studies a blood sample from an infected human individual was taken, cultured for several weeks and injected into mice and monoclonal antibodies were produced to explore which parasite antigens were detectable in those blood samples (Perrin *et al.*, 1981;Perrin *et al.*, 1982). These antibodies were then shown to inhibit *P. falciparum* parasite growth *in vitro* and conferred disease protection in *Saimiri* and *Aotus* monkeys (Perrin *et al.*, 1984;Perrin *et al.*, 1985;Siddiqui *et al.*, 1987). In the search for the major antigens involved in this protection a 140 kDa protein was detected. This protein was named serine rich protein I (SERPI) and later SERA5, as more members of the multigene family were identified (Knapp *et al.*, 1989). Protective immunity has now been established to be conferred by antibodies against a heterodimer comprised of the N- and C-terminus of SERA5. Since the early findings SERAs have been investigated as drug and vaccine targets. Nine *sera* genes were identified within the *P. falciparum* genome, with *sera1-8* located on chromosome 2 in a tandem head-to-tail array (with *sera4*, *5* and *6* in the centre) and *sera9* located on chromosome 9 (Gardner *et al.*, 2002). In *P. falciparum* all *seras*

encoded on chromosome two are clustered between a conserved hypothetical protein and a putative iron-sulfur assembly protein gene (hesB; Gardner *et al.*, 2002). All of these *seras* except *sera8* are transcribed in the asexual blood stage forms, though not all have been detected at the protein level (e.g. *sera1* and *sera2*). Expression levels of the *seras* vary extensively, with *sera4*, *sera5* and *sera6* the most abundantly transcribed (Aoki *et al.*, 2002;Miller *et al.*, 2002b). The SERA nomenclature is based on each gene's position on the chromosome in relation to other *seras*. Therefore SERA orthologues throughout all *Plasmodium* species do not have the same names (see Figure 10). All the *P. falciparum* SERAs contain a central papain-like putative catalytic domain with a catalytic triad present in all nine (Hodder *et al.*, 2003). Sequence identity within the *P. falciparum* SERA family is remarkably high and most variation can be found within the N-terminus (see Figure 11). The best example here is SERA5 with an extensive stretch of up to 37 Ser residues in the N-terminus which none of the other family members encode for. *P. falciparum* SERA1-5 and SERA9 are "serine-type" SERAs (SERAser) and form their own monophyletic group characterised by a substitution of the catalytic site Cys residue with a Ser residue (Kiefer *et al.*, 1996;Hodder *et al.*, 2003;Bourgon *et al.*, 2004). Despite this substitution, recombinant SERA5 was shown to possess weak catalytic activity (Hodder *et al.*, 2003). However, it was previously shown that a substitution of the active site cysteine residue with a serine residue in other papain-like proteases (e.g. cathepsin L) interferes with enzymatic activity (Clark *et al.*, 1978;Coulombe *et al.*, 1996). On the other hand, the silicateins, which are expressed by the sponge *Tethya aurantia*, also show a Cys to Ser substitution like the SERAsers but are fully functional active hydrolases (Shimizu *et al.*, 1998;Muller *et al.*, 2007). SERA6, 7 and 8 are "cysteine-type" SERAs (SERAcys), containing a canonical Cys residue in the predicted catalytic triad. SERA8 is not expressed in asexual stages and can only be found in mosquito stages of *P. falciparum*. SERA5 is the most abundant member of the SERA family in *P. falciparum* and is located in the PV (Delplace *et al.*, 1987;Debrabant *et al.*, 1992). This is of interest as *P. falciparum* proteases have been shown to be highly abundant in the PV (Nyalwidhe *et al.*, 2006). The refractoriness to genetic disruption of *sera5* and *sera6* (Miller *et al.*, 2002b;McCoubrie *et al.*, 2007) and also the inhibition of *P. falciparum* multiplication and growth by antibodies against SERA5 suggests that they play a major role in the parasite life-cycle (Pang *et al.*, 1999). Additionally SERA5 has been shown to undergo extensive proteolytic processing just moments before egress. The protease responsible for the majority of this processing is PfSUB1, a key player in egress (see 1.5.3.2). The most convincing evidence of the importance of the SERAs derives from a beautiful study by Aly and Matuschewski. The *P. berghei* orthologue of SERA8, *ecp 1*, was disrupted by single cross-over homologous recombination leading to two truncated

versions of the protein. The knockout had no effect on parasite growth in the asexual blood stages. The faulty ECP1 expression however affected the sexual *P. berghei* mosquito stages. Fully developed and viable sporozoites were incapable of rupturing the oocyst wall (Aly *et al.*, 2005). Aly and Matuschewski proposed the involvement of ECP1 in the degradation of the inner oocyst wall as *ecp1* knockouts showed an altered proteolytic processing pattern of the circumsporozoite protein (CSP), a highly abundant protein located on the surface of sporozoites. Although ECP1 has yet to be shown to have protease activity and the oocyst structure differs from that of the RBC, an important role in parasite egress was postulated for the putative SERA enzymes as orthologues of ECP1 (Dowse *et al.*, 2008).

In summary the SERA family members are most important from two different perspectives. First, as vaccine targets as they confer anti-disease immunity which is due to the postulated role of the most prominent family member SERA5 in invasion. Second, based on evidence derived from genetic manipulation and proteolytic processing studies they are of major importance for the release of parasites from their host cells and possibly for subsequent invasion.

1.6.2 SERA5: a “serine-type” SERA

SERA5 is one of the most abundant proteins in the *P. falciparum* trophozoite and schizont stage (Lasonder *et al.*, 2002), reinforcing interest in this protein. SERA5 was originally named P140, P126, P113 and SERP I and is by far the best studied and understood member of the SERA family. A first appearance in the spotlight can be dated back to the early 1980s. In the first studies, monkeys immunised with a 140 kDa molecule were partially protected against *P. falciparum* infection *in vivo* (Perrin *et al.*, 1981;Perrin *et al.*, 1984). More recent data suggests that human plasma in endemic areas show high anti-SERA5 antibody titers (Okech *et al.*, 2001;Okech *et al.*, 2006). This is supported by earlier findings that SERA5 fragments induce protection against challenge infection (Inselburg *et al.*, 1991). SERA5 antibodies not only confer anti-disease immunity but also protect against egress or invasion *in vitro* (Chulay *et al.*, 1987;Fox *et al.*, 1997;Pang *et al.*, 1998;Pang *et al.*, 1999). In addition Delplace *et al.* observed that a 126 kDa protein which is expressed in late schizont stages disappears and a 50 kDa fragment appears in culture supernatant of egressed *P. falciparum* merozoites *in vitro* (Delplace *et al.*, 1985;Delplace *et al.*, 1987;Delplace *et al.*, 1988). This fragment was later shown to be SERA5. Discrepancies in the molecular weight derive from the observation that the molecular weight of SERA5 is predicted as 126 kDa but it runs at around 130 kDa on an SDS-PAGE. SERA5 is expressed as a precursor protein and undergoes several proteolytic processing steps (Li *et al.*, 2002a;Yeoh *et al.*, 2007). The 126 kDa precursor protein (a putative inactive zymogen)

is cleaved in a first event into two fragments of 47 kDa and 73 kDa (see Figure 12, SERA5st1). The latter is further processed near its C-terminal side releasing a 56 kDa and an 18 kDa fragment (see Figure 12, SERA5st2). Both events have been shown to be mediated by PfSUB1 (Yeoh *et al.*, 2007;Arastu-Kapur *et al.*, 2008). The 47 kDa region undergoes further cleavage to two 25 kDa fragments (see Figure 12, SERA5st3) although this third processing step appears to be allele-specific (Li *et al.*, 2002b). The 56 kDa fragment containing the putative catalytic domain is further processed near the C-terminus to a 50 kDa fragment, still carrying the catalytic domain, and a 6 kDa fragment (see Figure 12, SERA5stX). Proteolytic processing of SERA5 was shown to only be induced under naturally mediated egress. Manual and mechanically induced rupture of *P. falciparum* cultures only releases the precursor into the culture supernatant. This suggests that SERA5 processing does not occur randomly due to unspecific protease activity but is more likely to be a targeted and highly regulated step.

The molecule mediating the 56 kDa to 50 kDa conversion event was identified as a leupeptin and E64 sensitive cysteine protease (Li *et al.*, 2002a) and called Protease X (Michael Blackman, unpublished). The SERA5 50 kDa fragment is highly abundant in culture supernatant just after merozoite release whereas it cannot be detected in late schizonts. This suggests that the last SERA5 processing step coincides with schizont rupture (Yeoh *et al.*, 2007). PfSUB1 processing sites are conserved within the other asexual stage *P. falciparum* SERAs (see Figure 13; Yeoh *et al.*, 2007). Although SERA5 might be the most abundant protein of the SERA family (Aoki *et al.*, 2002), recent data suggests that gene knockouts of the *P. berghei* orthologue of SERA5 (PbSERA2) have no effect on the parasite life-cycle but induce the up-regulation of the SERA6 orthologue (PbSERA3;Putrianti *et al.*, 2009). There is also very recent evidence that SERA5 is an essential protein but is not an essential protease; Stallmach *et al.* have replaced the active site Ser with an Ala residue in *P. falciparum* cultures with no observable phenotype but were unsuccessful in deleting the entire *sera5* gene (Robert Stallmach, unpublished). The question of what is the role of SERA5 remains, as it appears to be essential and is one of the most abundant proteins of *P. falciparum*.

1.7 SERA6, a putative malarial cysteine protease

SERA6 is the third most abundant of the SERA family members in the trophozoite and schizont parasite stages of *P. falciparum* (Lasonder *et al.*, 2002). It was first described as SERP H by Knapp *et al.* in 1991 (Knapp *et al.*, 1991) as a homologue of SERA5. Based on evidence derived from immuno EM, IFAs and a proteomic approach, SERA6 appears to be localised to the PV (Knapp *et al.*, 1991;Aoki *et al.*,

2002;Nyalwidhe *et al.*, 2006). However, low-resolution immuno EM figures, unspecific IFA labelling and a low peptide count from proteomic approaches mean that further verification is required in order to convincingly confirm the exact subcellular localisation of SERA6. Expression levels peak late in the *P. falciparum* asexual life stages.

SERA6 comprises 1031 amino acids and is encoded by 3096 bp over four exons (PlasmoDB: http://plasmodb.org/plasmo/) with a predicted N-terminal signal peptide targeting the protein into the secretory pathway (accession number PFB0335). It belongs to group III of the cysteine type SERAs (Arisue *et al.*, 2011). The predicted molecular weight of the full-length SERA6 is around 120 kDa. The entire SERA6 protein contains 29 cysteine residues. The best characterised SERA6 orthologue is *P. berghei* SERA3 and both SERA6 and PbSERA3 remain refractory to gene disruption (Miller *et al.*, 2002b;Kooij *et al.*, 2005;Arisue *et al.*, 2007;Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009).

The putative catalytic domain was identified via alignments with SERA5, *P. falciparum* SERAs and papain and includes the papain-like protease family canonical Cys, His and Asn residues within its predicted active site. Although they have one major discrepancy in the putative catalytic site, SERA5 and SERA6 exhibit an overall sequence identity of 40% (Baker *et al.*, 2001). SERA6 processing has yet to be fully described although unpublished data by Aoki *et al.* supports the extensive cleavage of SERA4 and SERA6 (McCoubrie *et al.*, 2007). Yeoh *et al.* further showed sequence conservation of the PfSUB1 cleavage sites across most *P. falciparum* SERAs, and suggested that SERA6 undergoes cleavage similar to SERA5 processing, however not necessarily in the same order and into different sized fragments (see Figure 14). In the latter study processing of SERA6 by recombinant PfSUB1 was proposed. Unfortunately antibodies used in this study were of poor specificity and strongly cross-reacted with other prominent SERA family members such as SERA4. Alignments with all *P. falciparum* SERAs and the known PfSUB1 processing sites suggest that SERA6 undergoes processing by PfSUB1 at two distinct sites (see Figure 14, SERA6st1 and SERA6st2), releasing a central fragment carrying the papain-like domain. PfSUB1 processing possibly instigates a conversion of SERA6 to an active protease.

The putative catalytic region of SERA6 has a high degree of identity with the catalytic region of papain, and modelling of SERA6 based on the falcipains suggests that SERA6 should be an active cysteine protease. Together with the likely subcellular localisation of SERA6 to the PV, and the effect of cysteine protease inhibitors on egress, this strongly suggests that SERA6 indeed has all the attributes which are required for a cysteine protease activity with a possible role in egress.

1.7.1 SERA6 and its heterodimer of terminal fragments

The N- and C-terminus of SERA5 remain connected via disulfide-bridges and translocate to the merozoite surface where they play an important role in merozoite invasion. The fate of the N- and C-terminus of SERA6 is unknown. Proteomic data show some SERA6 peptide counts from merozoite derived material (PlasmoDB: http://plasmodb.org/plasmo/). No experimental evidence of the translocation of the disulfide-linked termini to the merozoite surface has been established.

1.7.2 SERA6 and its orthologues

As described above, SERAs in 18 *Plasmodium* species have now been identified (Arisue *et al.*, 2011). All of these species carry a *sera6* orthologue in their genome with a high degree of identity (see Table 2). The overall high level of conservation between SERA6 and its orthologues possibly indicates an important conserved role for those putative cysteine proteases. This is supported by the fact that the variation in SERA numbers for *Plasmodium* species is due to group IV, the serine type SERAs. All the cysteine type SERAs are highly conserved and all known species encode one copy of groups I-III (see 1.6). The best studied SERA6 orthologue is *P. berghei* SERA3.

1.7.3 *P. berghei* SERA3: A SERA6 model

P. berghei orthologues of the *P. falciparum* SERAs have recently been identified followed by thorough investigation of the orthologues of SERA4, SERA5 and SERA6 (Arisue *et al.*, 2007). PbSERA3 has been shown to be expressed very late in the hepatocyte stage of *P. berghei* and is located to the PV (Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009). Processing of the putative protease moments before the liberation of merozoites has been suggested (Schmidt-Christensen *et al.*, 2008). Interestingly PbSERA3 is refractory to genetic manipulation in *in vitro* cultures and the disruption of the SERA4 and SERA5 orthologues either separately or combined induces upregulation of PbSERA3 (Putrianti *et al.*, 2009). The orthologues of *P. falciparum* SERA4 and SERA5 are expressed in blood stages. PbSERA3 mRNA could also be detected in asexual blood stage infections in the rodent host, although it undergoes strong upregulation in late hepatocyte stages (Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009).

SERA6 and PbSERA3 share an overall identity of 54%. The papain-like protease domain is highly conserved with 67% identity (see Figure 15). Alignments with SERA5 and SERA6 reveal that PbSERA3 indeed has cleavage sites which are similar to the PfSUB1 recognition motif. Substrate peptides based on those sites were shown to be proteolytically processed by recombinant *P. berghei* SUB1 (Catherine Suarez,

unpublished). High sequence identity and possibly analogous processing patterns make PbSERA3 an intriguing model to improve the understanding of the biological role of SERA6 in the parasite life-cycle.

1.8 Hypothesis and Aims

1.8.1 Summary

The active rupture of iRBCs by the human malaria parasite *P. falciparum* is an essential step in the parasite life-cycle, preceding re-invasion of non-infected RBCs. Convincing evidence shows that proteases are key players in this step, called egress, and therefore could serve as excellent targets for the development of new antimalarials.

Serine repeat antigens (SERAs) are a family of proteins expressed in the asexual blood stage forms (Miller *et al.*, 2002b) of *Plasmodium spp.* and all contain a central papain-like domain which identifies them as putative cysteine proteases. *P. falciparum* SERA6 is one of the most abundant SERAs in late parasite stages and its gene is refractory to disruption, suggesting an essential role for SERA6 in parasite survival. The timing of SERA6 expression and its putative localisation to the PV, a parasite induced vacuole in which the parasite develops and multiplies, make SERA6 an intriguing target for further investigation.

1.8.2 Working hypothesis

Recent evidence and a variety of well conducted approaches have delivered very convincing evidence to postulate that SERA6 is a cysteine protease that contributes to a proteolytic cascade leading to malaria parasite egress through rupture of the PVM or the EPM.

1.8.3 Main questions to address in this work

The main aims of this study were firstly to verify previous findings of SERA6 localisation, confirm its putative processing by PfSUB1, identify any enzymatic activity and investigate its role in parasite egress. Briefly, this work will explore: (A) SERA6 localisation, (B) SERA6 processing, and (C) SERA6 proteolytic activity. Questions to be addressed are as follows:

A) Is SERA6 a soluble PV protein?

B) Is SERA6 a substrate of PfSUB1?
    I. Is SERA6 processing by PfSUB1 indispensable for parasite survival?

C) Is SERA6 a cysteine protease?

I. Is the cleavage by PfSUB1 an activation event?
II. Is SERA6 activity indispensable for parasite survival?

A) SERA6: a soluble PV protein?

It was first aimed to express recombinant SERA6 to provide a tool of utmost importance for further analysis, both for antibody production and for activity studies. Previously-produced SERA6 antibodies were non-specific and showed strong cross-reactivity with other members of the SERA family (Yeoh *et al.*, 2007). Several different recombinant SERA6 fragments corresponding to the N-terminus, C-terminus and the central domain were aimed to be expressed in *E. coli* and COS-7 cells. The fragments were as follows: P64 fragment (homologue to SERA5 P56), designated as SERA6 Central1 (S6C1), P60 (homologue to SERA5 P50) fragments, termed SERA6 Central 2 (S6C2), and the full-length SERA6 (S6-FL). Some of these fragments were later used in PfSUB1 processing studies. In order to confirm previous studies we attempted to localise SERA6 by analysing parasite cultures with two different approaches. Firstly, by IFA with the newly generated antibodies and colocalisation with markers for various subcellular compartments and organelles. Secondly, by biochemical fractionation of the parasite which separates the contents of the PV from contents of the RBC and parasite cytosol, followed by immunochemical detection.

B) Is SERA6 a substrate of PfSUB1?

SERA5 was previously shown to be cleaved by PfSUB1 at two sites and one additional allele specific site. Alignment of the SERA amino acid sequences (Yeoh *et al.*, 2007) revealed that other SERAs possess similar putative cleavage sites. Consequently a putative cleavage of SERA6 by PfSUB1 can be proposed. This processing is intriguing from several aspects. Firstly, this processing possibly occurs within the PV only moments before egress (like SERA5) which may indicate a putative role for SERA6 in egress. Secondly, the processing by PfSUB1 might lead to the activation of SERA6 which in return hydrolyses its substrate in order for the parasite to escape from the RBC.

To identify PfSUB1 processing sites in SERA6 it was aimed to employ rS6-FL (see above) in PfSUB1 processing assays. N-terminal sequencing would reveal the exact PfSUB1 processing sites in SERA6. Mutations would then be introduced at these processing sites in the recombinant SERA6, to specifically prevent cleavage by PfSUB1. To determine whether SERA6 is processed by PfSUB1 in the parasite, it was attempted to purify the endogenous protein from *P. falciparum* cultures and subjected to recombinant PfSUB1 processing assays.

I. Is SERA6 processing by PfSUB1 indispensable for parasite survival?

It was tried to replace the predicted PfSUB1 processing sites in *in vitro* parasite cultures to inhibit processing and analyse the effect on possible phenotypes. Cleavage sites were altered by mutations of P2 and P1 residues or P4-P1 deletion designed to block PfSUB1 processing. Mutations were introduced for each site separately or in combination. SERA5 is known to undergo a final processing step by a leupeptin sensitive cysteine protease, called Protease X. Based on alignments it was proposed that SERA6 might undergo similar processing as it encodes for such a processing site. Additionally the Protease X processing site in the parasite was replaced in order to investigate the role of that processing step.

C) Is SERA6 a cysteine protease activated by PfSUB1?

I. Is the cleavage by PfSUB1 an activation event?

Based on alignments with SERA5 it was predicted that SERA6 might be processed by PfSUB1 in order to be activated. To establish whether SERA6 is a cysteine protease it was attempted to express recombinant SERA6 either as the putatively activated fragment (S6C1, between predicted SERA6st1 and SERA6st2) or S6-FL. Recombinant expression of a putative active fragment included refolding of proteins derived from insoluble fractions and extensive testing of expression conditions and variation of expression hosts. Full-length SERA6 was subjected to cleavage assays with PfSUB1. This also aimed to verify whether PfSUB1 processing might be a SERA6 activation event leading to an active cysteine protease. SERA6 activity was tested on fluorogenic peptides, RBC ghosts, iRBC ghosts and zymograms with recombinant and parasite derived SERA6. As there was no specific inhibitor available controls were obtained by replacing the catalytic Cys with Ala in recombinantly derived SERA6.

Additionally it was attempted to recombinantly express PbSERA3, a SERA6 orthologue from the rodent malaria parasite *P. berghei*. The recombinant wild-type and Cys to Ala mutant were subjected to PfSUB1 activity assays and activity was assessed.

II. Is SERA6 activity indispensable for parasite survival?

In order to investigate SERA6 activity in the parasite it was attempted to replace the catalytic Cys with Ala in *P. falciparum* cultures by homologous recombination.

Table 1: Overview of all *seras* from 18 *Plasmodium* species

Summary of all known *sera* genes from 18 *Plasmodium* species. Number of *sera* genes, number of pseudogenes and the parasite's natural host are indicated. Data derived from Arisue *et al.*, 2007 and Arisue *et al.*, 2011.

Species	***No. sera genes***	***No. fragmented or pseudo sera genes***	***Host***
P. malariae	10	3	Humans
P. ovale	7	1	Humans
P. cynomolgi	11	2	Asian Old world monkeys
P. fieldi	9	2	Asian Old world monkeys
P. simiovale	9	2	Asian Old world monkeys
P. inui	7	4	Asian Old world monkeys
P. hylobati	7	0	Gibbons
P. coatneyi	7	1	Asian Old world monkeys
P. knowlesi	6	0	Humans Asian Old world monkeys
P. fragile	5	2	Asian Old world monkeys
P. gonderi	9	1	African Old world monkeys
P. vivax	12	1	Humans
P. berghei	5	0	Rodents
P. yoelii	5	0	Rodents
P. chabaudi	5	0	Rodents
P. falciparum	9	0	Humans
P. reichenowi (*sera* gene sequences not complete)	8	0	Chimpanzees
P. gallinaceum	3	0	Birds

Table 2: Overview of *P. falciparum* SERA6 orthologues

Species	***SERA6 orthologue nomenclature***	***Host***
P. malariae	PmaSERA8	Humans
P. ovale	PovSERA5	Humans
P. cynomolgi	PcySERA9	Asian Old world monkeys
P. fieldi	PfiSERA7	Asian Old world monkeys
P. simiovale	PsoSERA7	Asian Old world monkeys
P. inui	PinSERA5	Asian Old world monkeys
P. hylobati	PhySERA5	Gibbons
P. coatneyi	PcoSERA5	Asian Old world monkeys
P. knowlesi	PknSERA4	Humans Asian Old world monkeys
P. fragile	PfrSERA3	Asian Old world monkeys
P. gonderi	PgoSERA7	African Old world monkeys
P. vivax	PviSERA10	Humans
P. berghei	PbSERA3	Rodents
P. yoelii	PyoSERA3	Rodents
P. chabaudi	PchSERA3	Rodents
P. reichenowi (*sera* gene sequences not complete)	PreSERA6	Chimpanzees

Figure 1: Global Malaria Distribution in 2007

The map displays the global distribution of *P. falciparum* malaria in 2007. Low prevalence regions are marked in yellow and high endemic areas are marked in red. Countries with blue borders have reliable national health information system and very low *P. falciparum* infection rates Hay *et al.*, 2010.

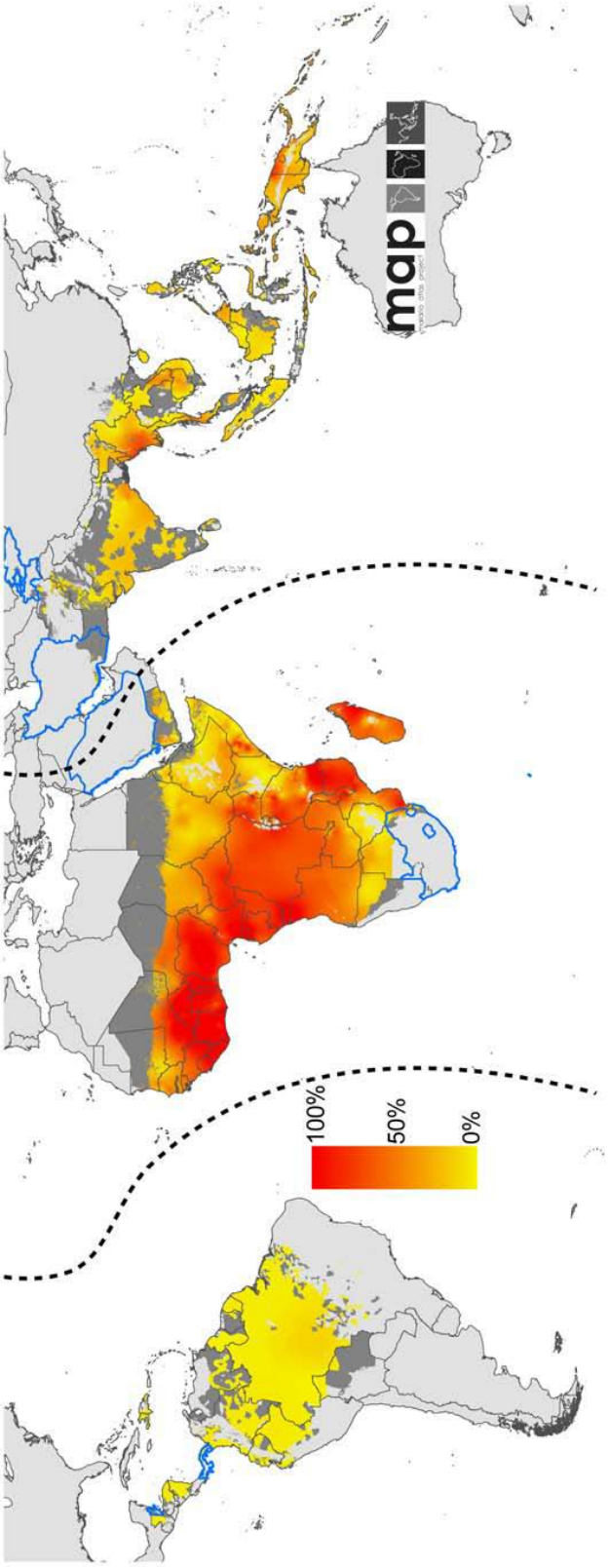

map
100%
50%
0%

Figure 2: Development of chloroquine resistance

Resistance to chloroquine originated in Southeast Asia. Chloroquine resistance first developed through a point mutation in the *pfcrt* gene which made the chloroquine mode of action ineffective. Areas of high prevalence of resistance are marked red, and numbers indicate the years in which resistance was first detected. Figure adapted from Wellems *et al.*, 2009.

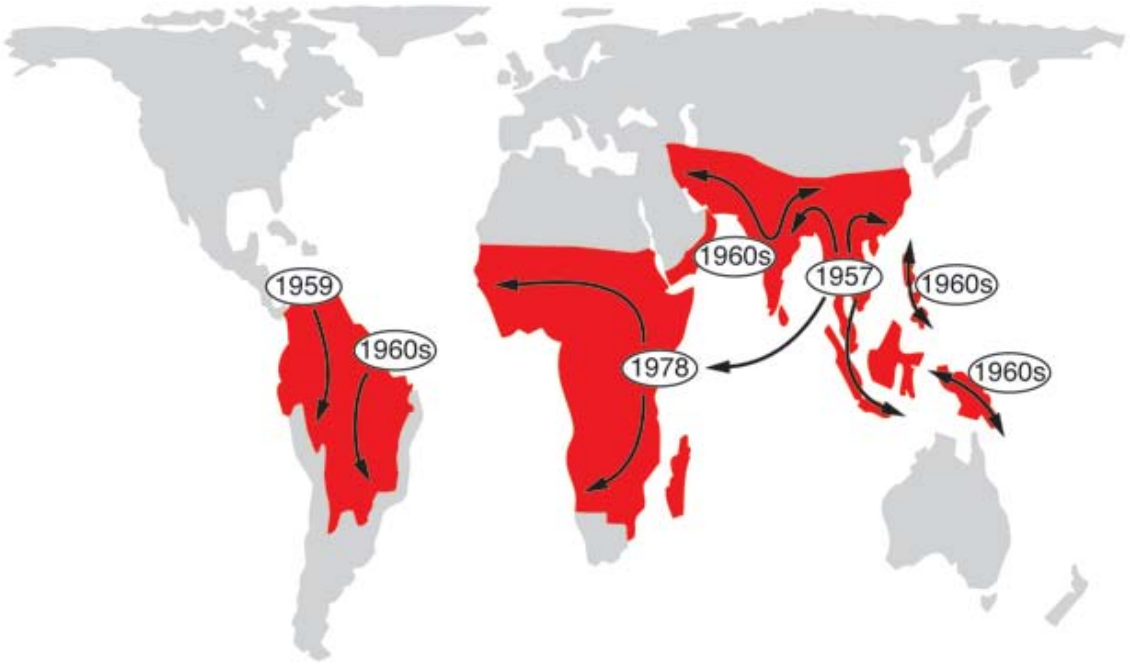

1959
1960s
1978
1960s
1957
1960s
1960s

Figure 3: Schematic of the *P. falciparum* life-cycle

The asexual cycle starts with the inoculation into human subcutaneous tissue of infective sporozoites by the female *Anopheles* vector. After the invasion of hepatocytes, parasites develop into exo-erythrocytic stage merozoites, followed by their release into the blood stream. These merozoites subsequently invade erythrocytes, initiating the intra-erythrocytic stage. This step is characterised by multiple cycles of RBC invasion, parasite growth and replication and discharge of fresh infective merozoites into the blood stream, ready to invade yet again. Triggered by a variety of conditions some of the parasites, however, follow a different path and develop into macro- and microgametocytes, which are essential for parasite transmission. Only those stages have the potential to begin gametogenesis in the mosquito after uptake by the dipteran vector in a blood meal. Fertilisation in the mosquito is followed by multiple rounds of asexual replication giving rise to new infective sporozoites that migrate to the mosquito's salivary glands, ready to be injected into a new human host, thus completing the parasite life-cycle. Figure adapted from Miller *et al.*, 2002a. RBC: red blood cell

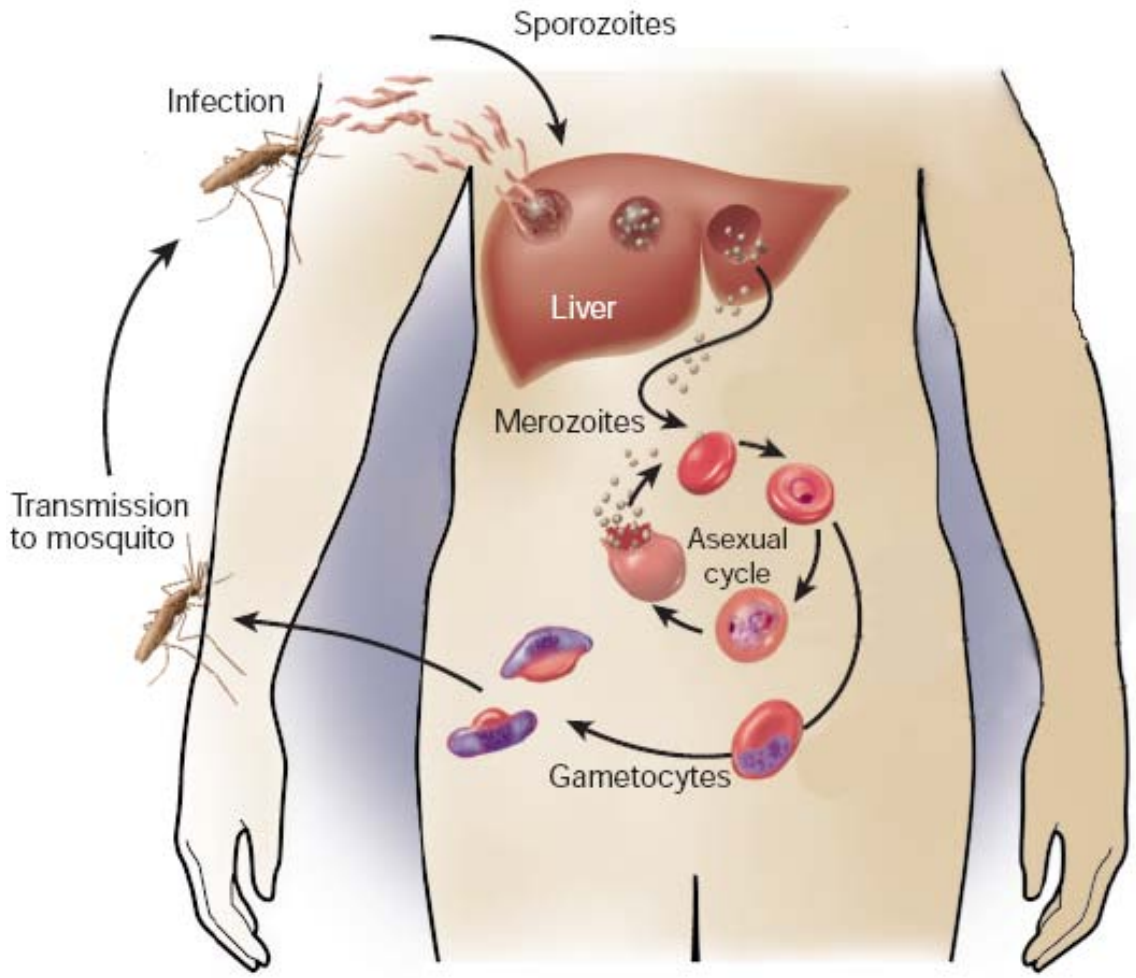

Sporozoites
Infection
Liver
Merozoites
Transmission
to mosquito
Asexual
cycle
Gametocytes

Figure 4: Schematic of *P. falciparum* intra-erythrocytic stage

Invasion of RBCs by merozoites. After the initial attachment and reorientation an electron-dense tight junction is formed between parasite and host cell and active invasion begins with the establishment of the PV surrounding the parasite (A/B). The parasite develops from ring (B) to trophozoite stage (C). The parasite matures and undergoes mitotic replication to form a schizont, containing new merozoites. Finally, merozoites are released from the iRBC into the blood stream (E). This step, like invasion, is a process actively governed by the parasite and termed egress. Figure adapted from Bannister *et al.*, 2003. RBC: red blood cell

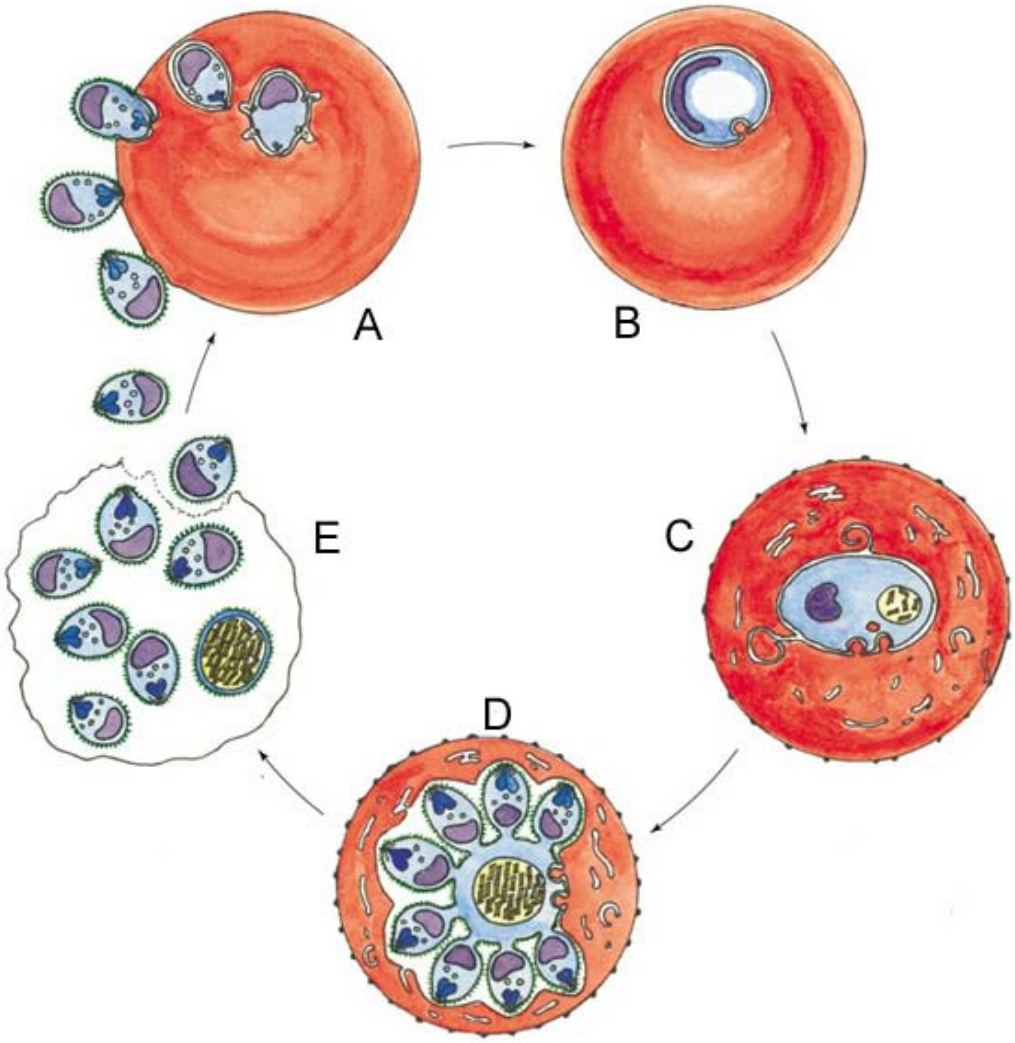

A
B
C
D
E

Figure 5: Ultrastructure of a *P. falciparum* merozoite

P. falciparum merozoites contain highly specialised organelles for egress and invasion of RBCs. Two to three microtubules run along the longitudinal axis. Proteins released from micronemes and club-shaped rhoptries are essential in invasion. Dense granules are heterogeneous in their contents and might perform different roles in the parasite life-cycle. All subsets of dense granules, plus exonemes which release PfSUB1 prior egress, play a role in host cell modification. The apical pole of the merozoite contains polar rings that are constituted of microtubules. The merozoite surface is coated with merozoite surface proteins (MSPs) which are involved in invasion of RBCs. The apicoplast is a red algae derived essential organelle with importance in parasite metabolism. Merozoites measure ~1 μm in diameter in the longitudinal axis. Figure adapted from Silmon de Monerri, 2010.

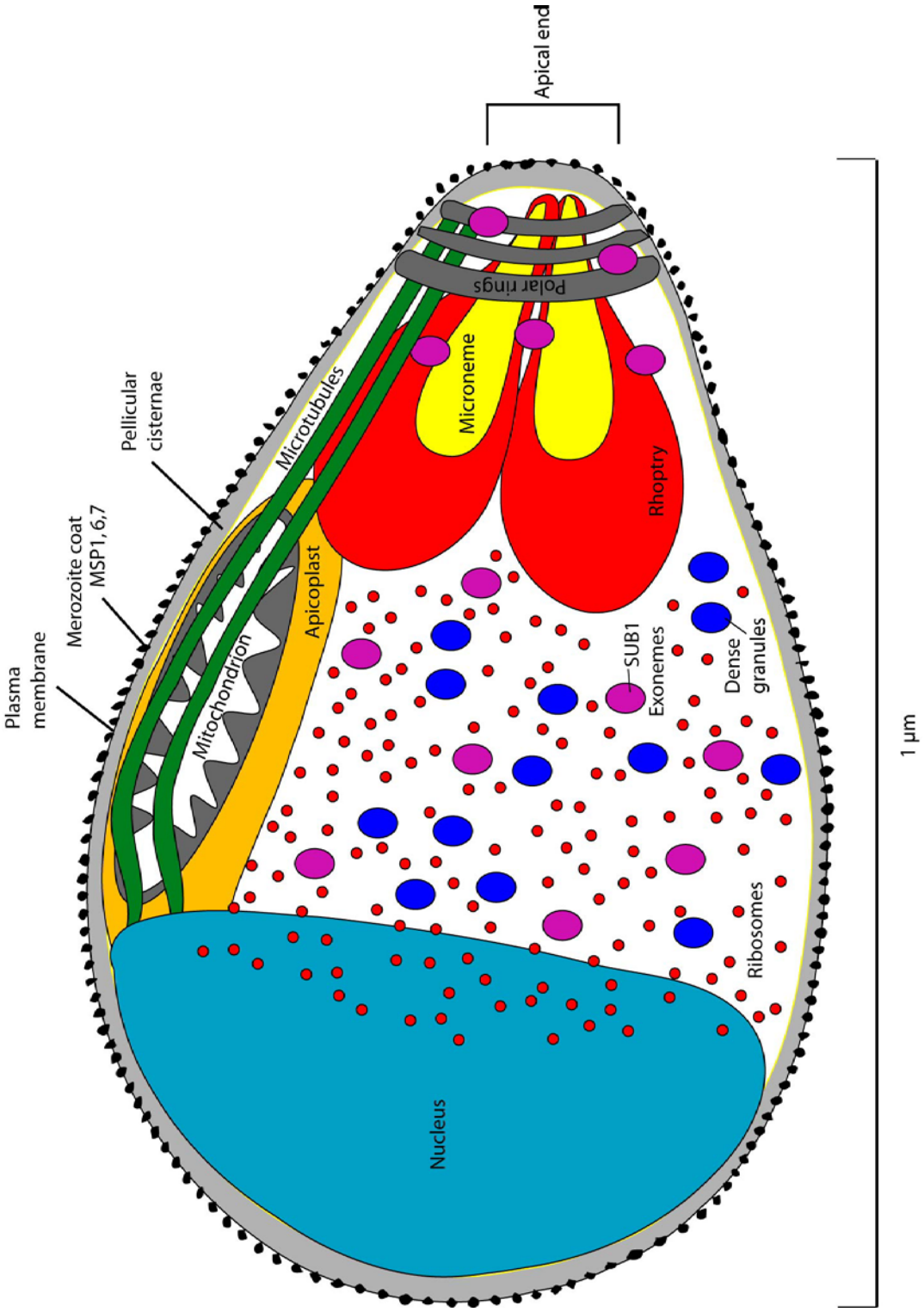

Apical end
Polar rings
Microneme
Rhoptry
Microtubules
Pellicular cisternae
Merozoite coat MSP1, 6,7
Plasma membrane
Mitochondrion
Apicoplast
SUB1
Exonemes
Dense granules
Ribosomes
Nucleus
1 µm

Figure 6: Invasion of erythrocytes by *P. falciparum* merozoites

The merozoite undergoes a primary weak attachment with the RBC which is mediated by surface receptors (A). The merozoite then re-orientates until the apical end reaches the RBC surface and the formation of a tight junction is established (B). These steps require highly regulated ligand-receptor interaction between the host cell and the merozoite. Translocation of tight junction components to the posterior pole of the merozoite pushes the merozoite into the RBC (C-E). Upon the start of the merozoite internalisation, merozoite surface adhesins are proteolytically released. The merozoite establishes a new compartment within the host cell (called the PV) during invasion. Figure adapted from Cowman *et al.*, 2006. RBC: red blood cell

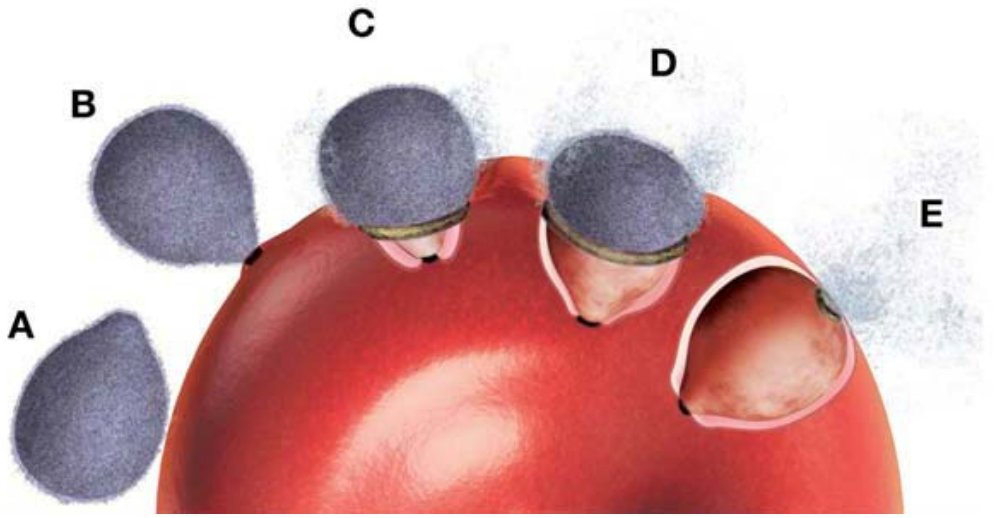

C
D
B
E
A

Figure 7: Detailed schematic of *P. falciparum* egress

Mature and newly formed merozoites are surrounded by still intact PVM (A). The PVM breaks down and viable merozoites move freely within the RBC cytosol (B). Swelling of the RBC (C) and degradation of the RBC skeleton (D) lead to the release of viable and invasion ready merozoites. Figure adapted from Cowman *et al.*, 2006.

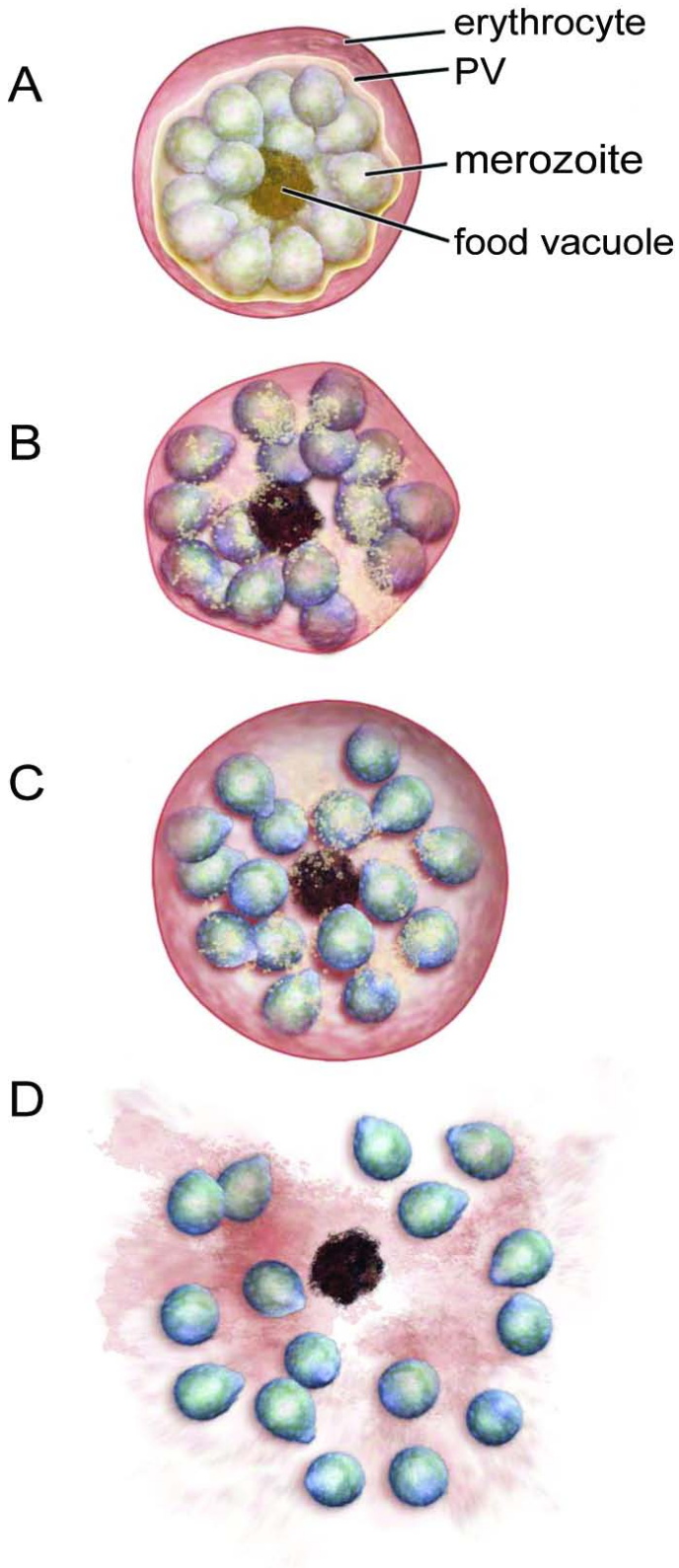

A
erythrocyte
PV
merozoite
food vacuole
B
C
D

Figure 8: Current models of parasite egress

A) Non-explosive egress via local fusion of the PVM with the EPM (Winograd *et al.*, 1999)

B) “Outside-in” model suggesting EPM rupture first and followed by PVM degradation as an exo-erythrocytic event (Salmon *et al.*, 2001;Baumeister *et al.*, 2006).

C) “Inside-out” model in which the PVM rupture prior to EPM disintegration (Wickham *et al.*, 2003)

D) Explosive RBC rupture with preceding “flower” formation of schizonts (Glushakova *et al.*, 2005). Figure adapted from Blackman, 2008.

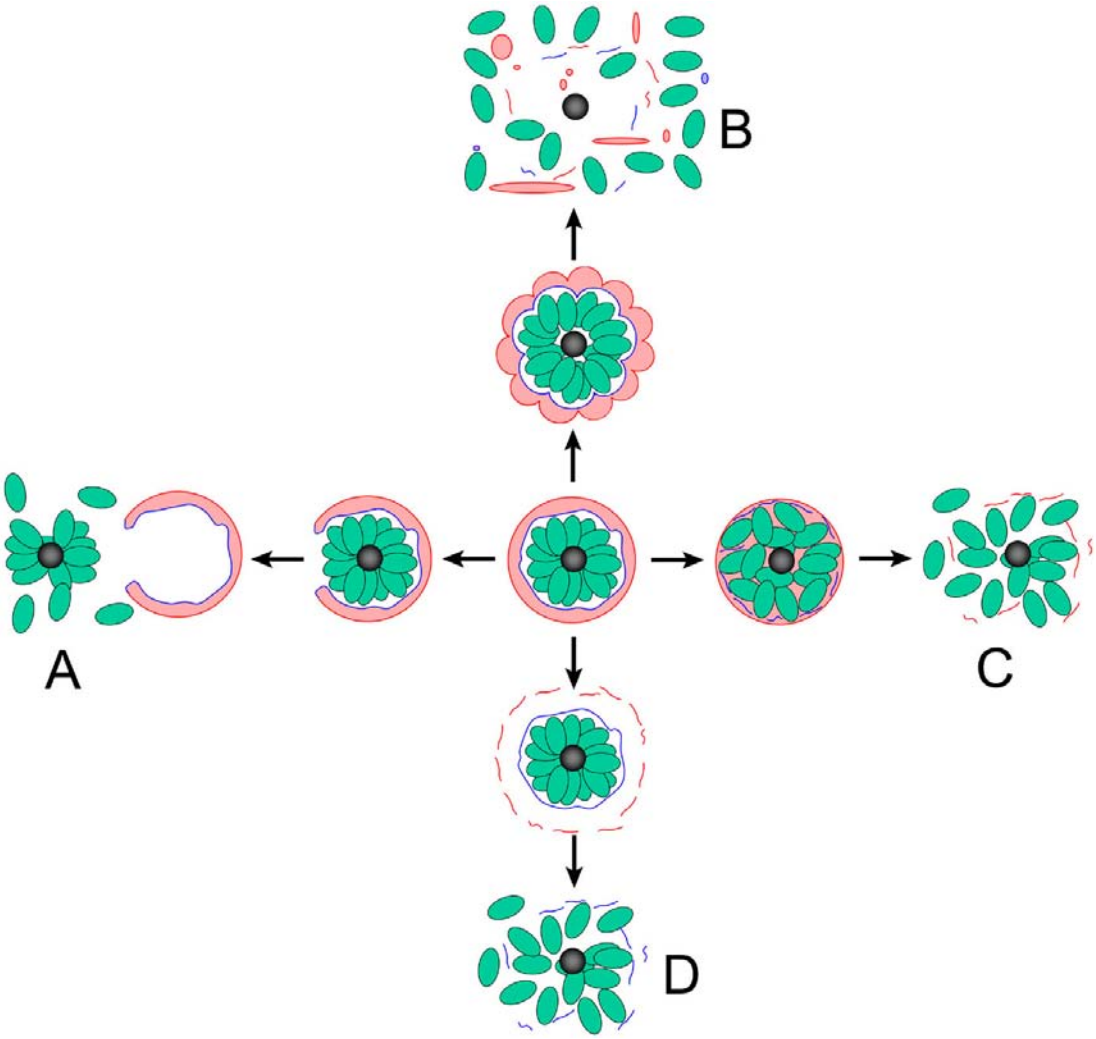

B
A
C
D

Figure 9: PfSUB1 substrate specificity

Graphical representation of a multiple sequence alignment of amino acid residues of known and predicted PfSUB1 cleavage sites (Silmon de Monerri, 2010;Silmon de Monerri *et al.*, 2011). The height of each residue stack represents the extent of the sequence conservation at that position. The height of each residue within that stack indicates the relative frequency of each residue. The arrow indicates the position of the scissile-bond. The colours of the amino acid residues are according to the chemical properties of the side-chains. Black, small: G,A; red, acidic: D,E; blue, basic: K,R,H; orange, aliphatic: L,V,I; green, uncharged polar: S,T,Y,N,Q; purple, non-polar, non-aliphatic: F,P.

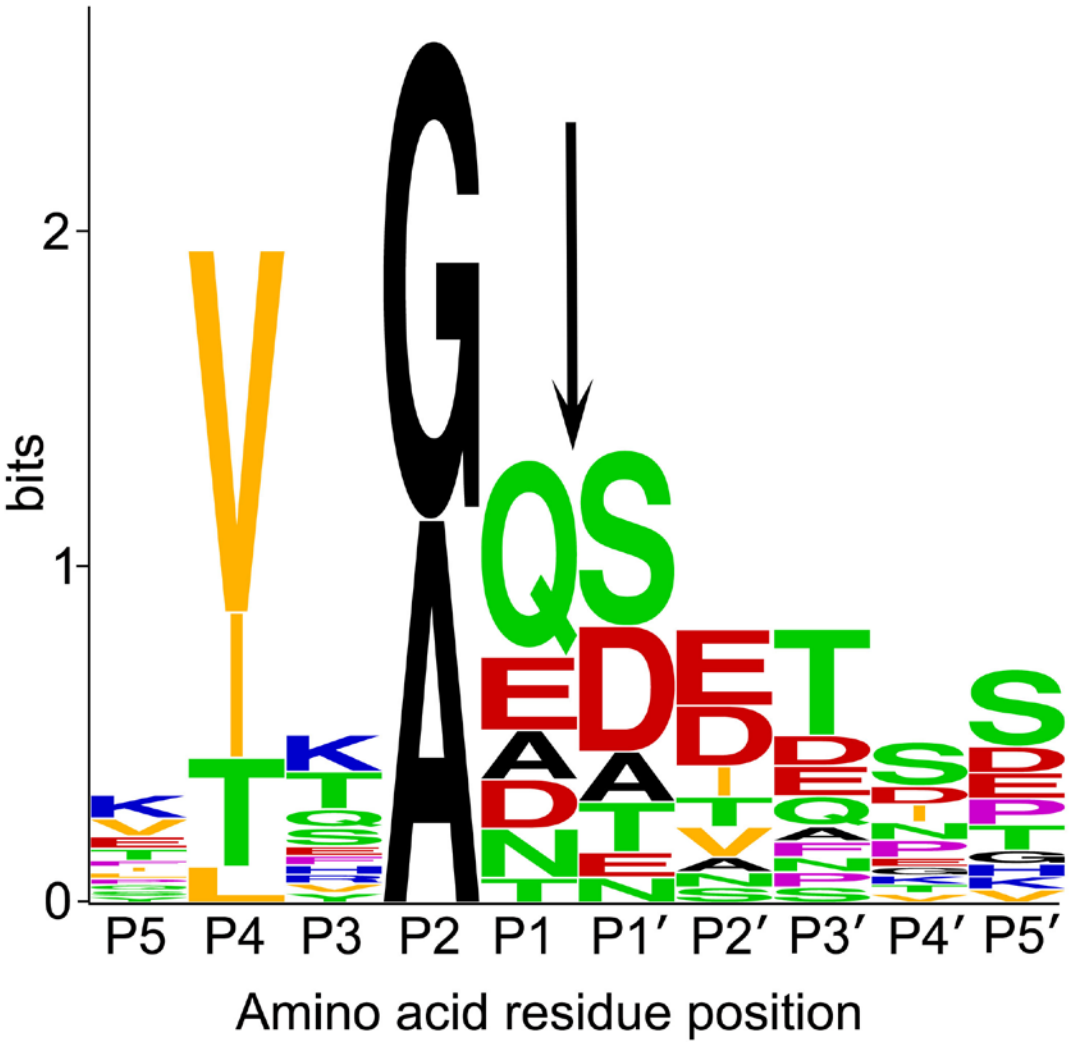

bits
2
1
0
P5
P4
P3
P2
P1
P1′
P2′
P3′
P4′
P5′
Amino acid residue position

Figure 10: Organisation of all known *sera* genes

Seras of 18 Plasmodium species are displayed according to their localisation on the chromosome and in relation to their orthologues. Each horizontal line represents one parasite species. All the genes are clustered between a conserved hypothetical protein gene and an iron-sulfur assembly protein gene, except *P. falciparum sera9*, *Plasmodium reichenowi sera9* and *P. gallinaceum sera3* which are located elsewhere. High conservation of *seras* from group I-III, cysteine-type SERAs, (green) is evident. Truncated genes and gene number variety can only be found in group IV, the serine-type SERAs (yellow). Figure adapted from Arisue *et al.*, 2011.

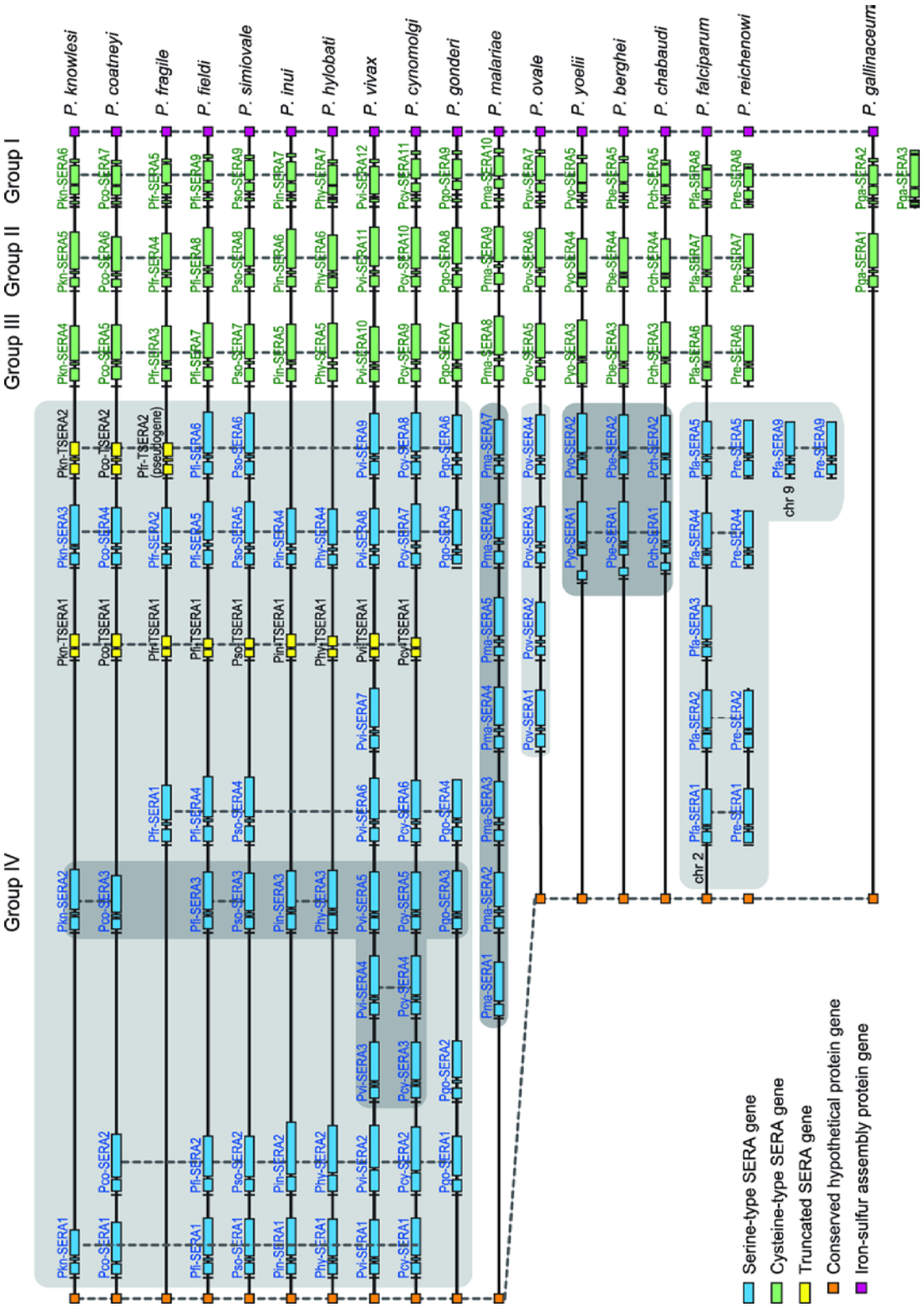

Group I
Group II
Group III
Group IV
P. knowlesi
P. coatneyi
P. fragile
P. fieldi
P. simiovale
P. inui
P. hylobati
P. vivax
P. cynomolgi
P. gonderi
P. malariae
P. ovale
P. yoelii
P. berghei
P. chabaudi
P. falciparum
P. reichenowi
P. gallinaceum
chr 2
chr 9
Serine-type SERA gene
Cysteine-type SERA gene
Truncated SERA gene
Conserved hypothetical protein gene
Iron-sulfur assembly protein gene

Figure 11: Conserved and non-conserved domains of *P. falciparum* SERAs

Consensus SERA structure for *P. falciparum*. Alignment was obtained with Clustal W. The overall structure shows the high degree of sequence conservation and homology between all *P. falciparum* SERAs (gray). All SERAs contain a considerable number of Cys residues, e.g. the *sera6* gene encodes 29 Cys, which appear to be positionally conserved (yellow bars). The central papain-like domain however, confers the highest degree of homology (light blue). Putative catalytic residues are indicated. Figure adapted from Yeoh *et al.*, 2007.

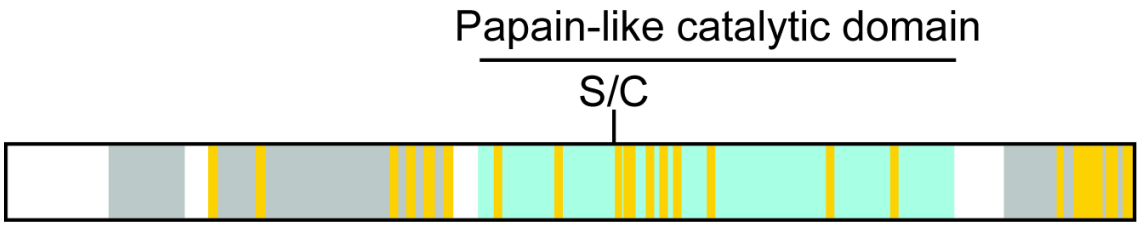

Papain-like catalytic domain
S/C

Figure 12: *P. falciparum* SERA5 undergoes extensive processing

SERA5 has been shown to be processed by PfSUB1 at two different sites, as well as at an allele-specific third cleavage site near the N-terminus in certain strains. In a first step PfSUB1 cleaves SERA5 in the N-terminal region (SERA5st1) generating two fragments, P47 and P73. A second cleavage near the C-terminus frees two fragments: P56, which carries the putative catalytic domain, and P18 at the very C-terminus (SERA5st2). P47 and P18 remain as a heterodimer, linked via a disulfide bond, and undergo translocation to the merozoite surface, where they are thought to play a role in invasion. The N-terminal fragment of the heterodimer undergoes an additional allele specific processing by PfSUB1 at a third site (SERA5st3), generating two fragments called P25 and P25'. Release of P56 has been suggested to act as an activation event for the protease. In the final processing step at a position just downstream of site 2, called P50C (SERA5stX), SERA5 is processed by an unknown protease which is leupeptin and E64 sensitive, releasing P50 and a C-terminal 6 kDa fragment. S-S indicates disulfide bonds; Amino acid sequences of the PfSUB1 cleavage sites are displayed. Presumed catalytic triad of papain-like domain are indicated. The final processing site was identified in the host lab (Sharon Yeoh and Michael Blackman, unpublished data).

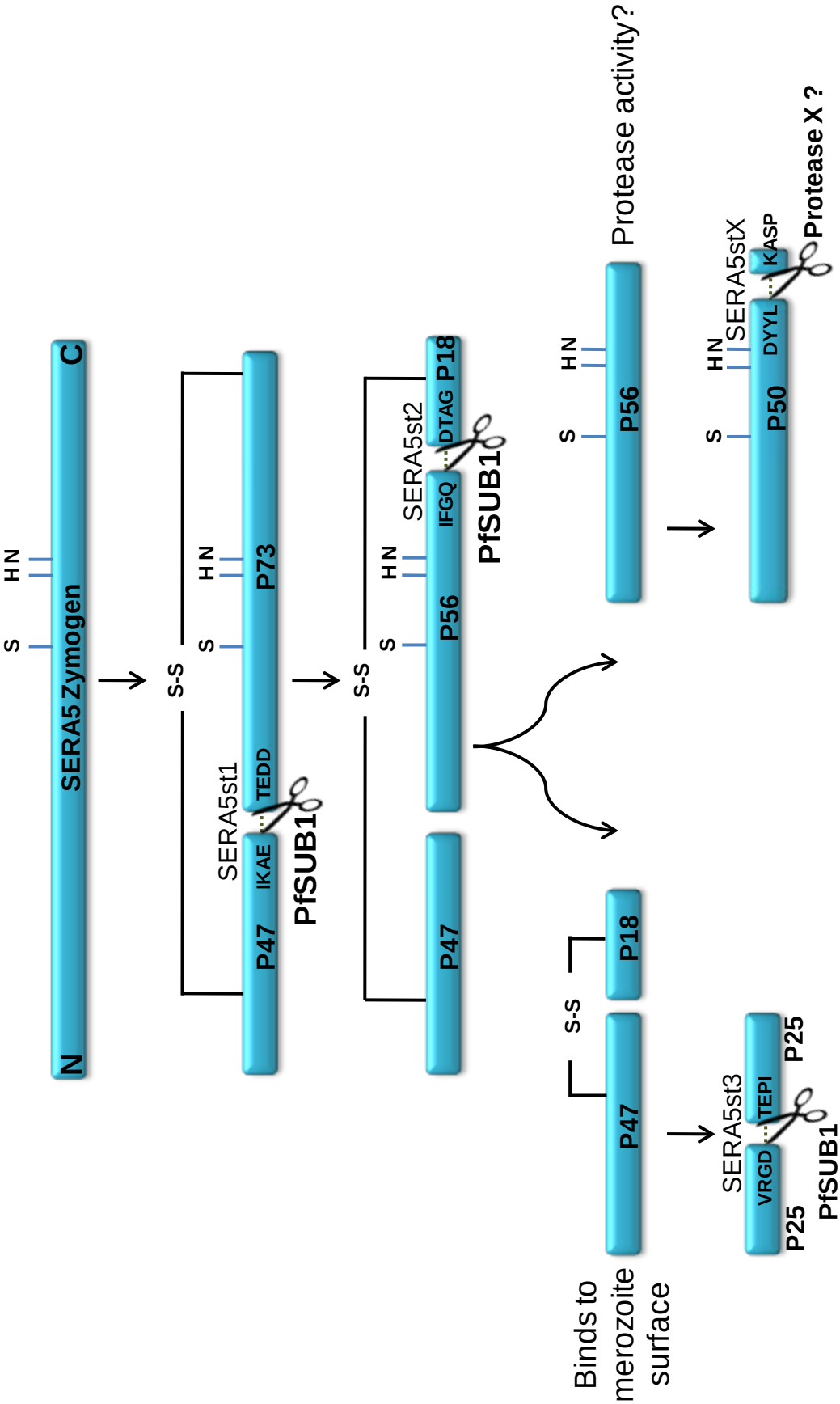

SERA5 Zymogen
N
C
H N
S
S-S
P73
P47
SERA5st1
IKAE
TEDD
PfSUB1
P56
P18
SERA5st2
IFGQ
DTAG
PfSUB1
Protease activity?
P56
P50
SERA5stX
DYYL
KASP
Protease X ?
P18
P47
Binds to merozoite surface
SERA5st3
TEPI
VRGD
P25
PfSUB1

Figure 13: Conservation of processing sites in *P. falciparum* SERAs

Based on alignments with known PfSUB1 processing sites in SERA5, other blood stage *P. falciparum* SERAs were shown to possess similar potential cleavage sites (Yeoh *et al.*, 2007). SERA5 is processed by PfSUB1 at three sites (SERA5st1, 2, 3) and by a second protease at one site (SERA5stX, Sharon Yeoh and Michael Blackman, unpublished data). Predicted processing sites for all other intra-erythrocytic SERAs are indicated below SERA5st1 and 2 (labelled with *st1 and st2*). Unpublished data suggests that SERA6 also encodes an additional PfSUB1-independent cysteine protease processing site (labelled with *stX*). Blue: Central papain-like domain.

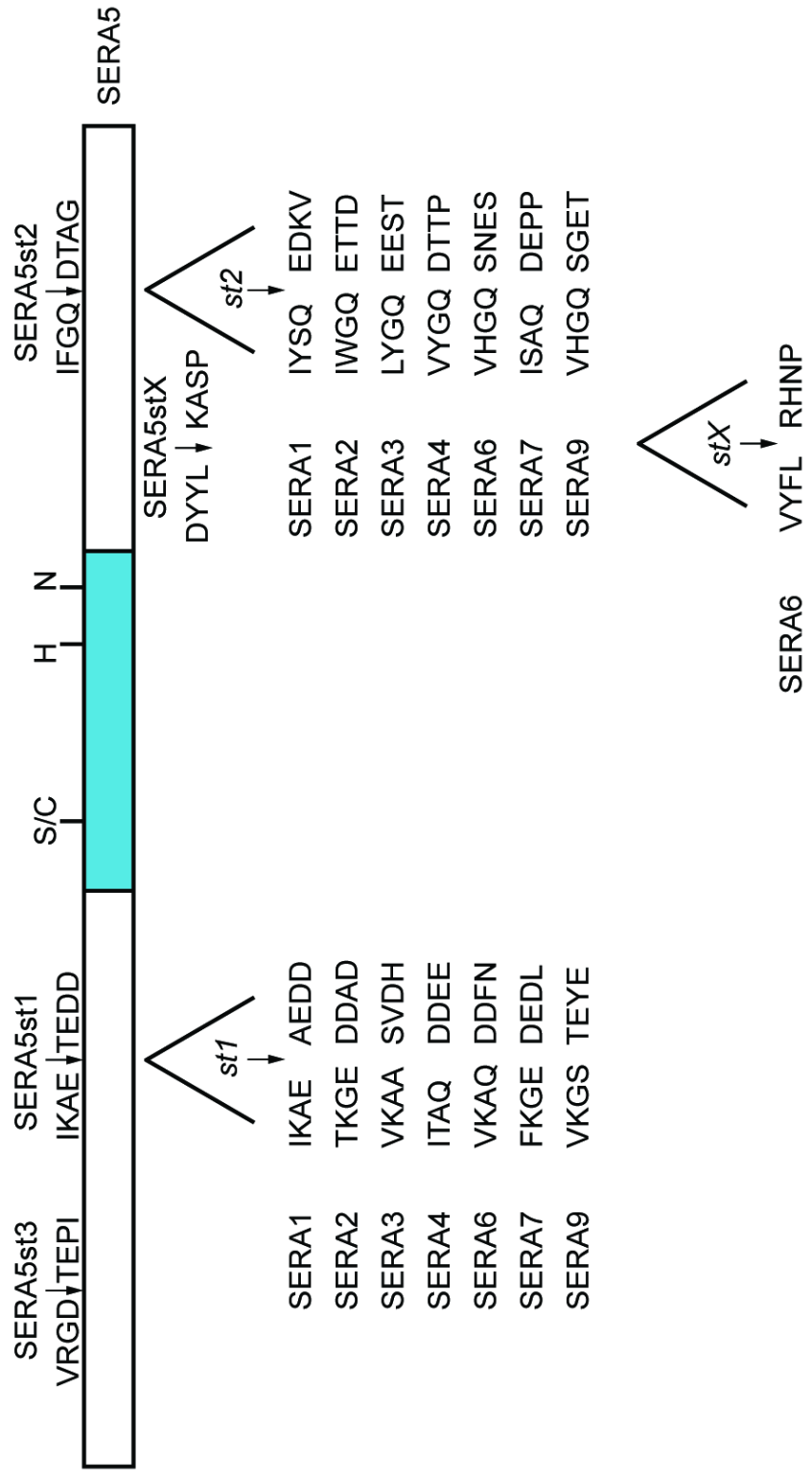

SERA5
SERA5st2
IFGQ DTAG
SERA5stX
DYYL KASP
st2
SERA1 IYSQ EDKV
SERA2 IWGQ ETTD
SERA3 LYGQ EEST
SERA4 VYGQ DTTP
SERA6 VHGQ SNES
SERA7 ISAQ DEPP
SERA9 VHGQ SGET
N
H
S/C
SERA5st1
IKAE TEDD
st1
SERA1 IKAE AEDD
SERA2 TKGE DDAD
SERA3 VKAA SVDH
SERA4 ITAQ DDEE
SERA6 VKAQ DDFN
SERA7 FKGE DEDL
SERA9 VKGS TEYE
SERA5st3
VRGD TEPI
stX
SERA6 VYFL RHNP

Figure 14: Predicted *P. falciparum* SERA6 processing by PfSUB1 and Protease X
PfSUB1 cleavage sites of SERA6 were predicted via the alignment of all intra-erythrocytic *P. falciparum* SERAs on the basis of the known SERA5 cleavage sites and the known PfSUB1 enzymatic binding pocket specificity (for sites SERA6st1 and SERA6st2). A first N-terminal cleavage site (SERA6st1) was predicted at the sequence VKAQ↓DDFN, releasing an N-terminal 38 kDa and C-terminal 76 kDa fragment. A second cleavage site (SERA6st2) is anticipated near the C-terminus at the amino acid sequence VHGQ↓SNES. Two fragments of 64 (S6C1) and 12 kDa would be generated. The 64 kDa fragment carries the putative catalytic domain and might have cysteine protease activity. An additional cleavage site (SERA6stX) for an unknown but leupeptin and E64 sensitive protease was modelled on the basis of the C-terminal SERA5 cleavage site (unpublished data). This final step possibly releases a 60 kDa fragment (S6C2), still carrying the putative catalytic domain, and a 4 kDa C-terminal fragment. A third N-terminal PfSUB1 site (such as is found for SERA5 in certain strains) is not predicted.

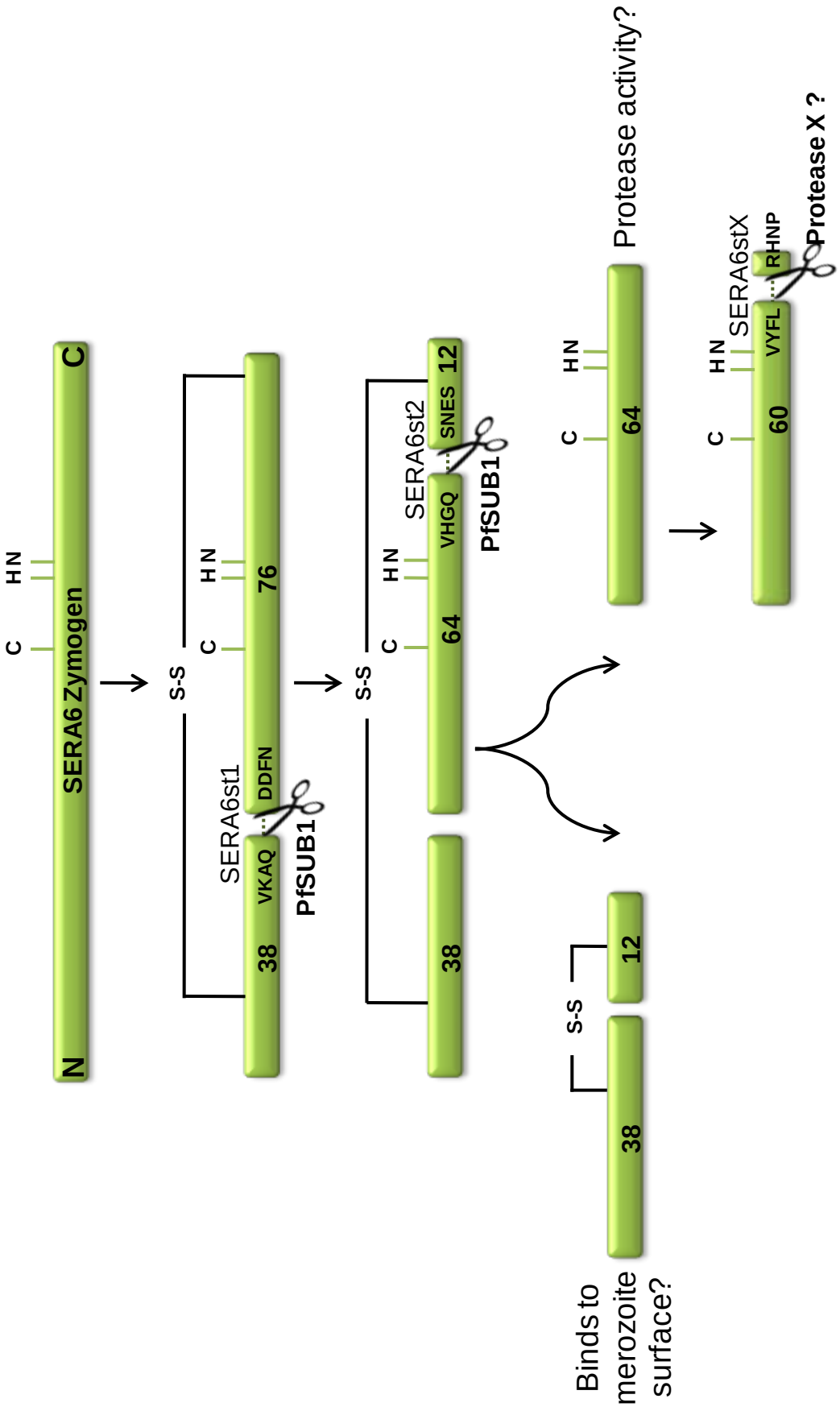

SERA6 Zymogen
N
C
H N
C
S-S
SERA6st1
PfSUB1
VKAQ
DDFN
38
76
SERA6st2
PfSUB1
VHGQ
SNES
64
12
Protease activity?
64
SERA6stX
Protease X ?
RHNP
VYFL
60
Binds to merozoite surface?
38
12

Figure 15: SERA6 and PbSERA3 are highly conserved

P. falciparum SERA6 and *P. berghei* SERA3 were aligned using Clustal W. Both SERAs share an overall identity of 54%. The central papain-like domains are 67% identical with the residues of the catalytic triad highly conserved (identical residues highlighted in blue, catalytic Cys residue in green). Putative PfSUB1 processing sites are conserved in both SERAs. The alignment suggests that PbSERA3 possibly undergoes a similar fate to SERA6 (red, arrow indicates scissile bond).

2 Materials and Methods

2.1 Parasite and COS-7 cultures

2.1.1 Maintenance and synchronisation of *P. falciparum* (Trager *et al.*, 1976;Lambros *et al.*, 1979;Trager, 1994)

Clone 3D7 of *P. falciparum* was cultured in plastic tissue flasks (Nunc™) in complete RPMI-1640 Albumax cell culture medium (Invitrogen™ GIBCO®), supplemented with 2 mM L-Glutamine. O^+ human RBCs were obtained from the National Blood Bank and iRBCs were cultured at 1-2% hematocrit. Parasite cultures were regularly synchronised and enriched via density gradient centrifugation with Percoll (GE Healthcare) selecting for iRBC harbouring late parasite stages (schizonts) to a parasitaemia >90% (Rivadeneira *et al.*, 1983;Dluzewski *et al.*, 1984). Infected RBCs were harvested by centrifugation for 3 minutes (min) at 1,000 x *g* then resuspended in 4 ml of RPMI (no Albumax) and gently layered on top of a warm 63% Percoll-PBS solution. For separation the Percoll-parasite layer mix was centrifuged for 11 minutes at 1,000 x *g*, then the top layer including the schizont iRBCs was removed and washed in Albumax free RPMI. The schizont pellet was resuspended in an appropriate volume of parasite cell culture medium supplemented by an appropriate volume of washed RBCs, transferred to sterile cell culture flasks, gassed with a 1% O_2, 3% CO_2, N_2 gas mixture and left for re-invasion for 2 hours (h) with gentle shaking. After invasion the cells were incubated with 5% sorbitol at 37°C for 8 min to lyse residual schizonts, then washed and recultured at 37°C. Parasitaemia was routinely assessed by thin blood films which were fixed with 100% methanol and stained with 10% Giemsa (VWR International) for 5-10 min. Parasite stages were identified via oil-immersion light microscopy.

2.1.2 Preparation of mature schizont extracts

Parasites were synchronised and highly mature schizonts were enriched as described in 2.1.1, then recultured and smeared every 20 min until first ring stages began to appear. The cultures were then pelleted at 1,000 x *g* for 3 min and stored in aliquots at -80°C until use.

2.1.3 Genomic DNA extraction

Genomic DNA from non-transfected and transfected parasite lines was obtained by lysis of a synchronous trophozoite culture with a parasitaemia of 5%. Parasites were pelleted, resuspended in 0.15% saponin (BDH Laboratory Supplies) in PBS, incubated at room temperature (RT) for 5 min and pelleted at 12,400 x *g* for 5 min. DNA was extracted according to manufacturer's instructions with a DNeasy Blood & Tissue Kit (Qiagen).

2.1.4 Continuous COS-7 culture

COS-7 cells were thawed quickly from liquid N_2 in a 37°C water bath and resuspended in 0.5 ml prewarmed Dulbecco's modified Eagle medium/GlutaMAX/Pyruvate/Glucose (Invitrogen™ GIBCO®) containing 10% fetal calf serum (FCS), 4 mM L-Glutamine and 0.5% Penicillin/Streptomycin 10,000 units (Pen/Strep; SIGMA), referred to as complete DMEM. Cells were then transferred to a 50 ml Falcon tube and a further 10 ml of complete DMEM was added before the mix was pelleted at 140 x *g* for 5 min. The supernatant was discarded, the pellet was resuspended in 40 ml complete DMEM and transferred to a 175 cm^2 culture flask (Nunc™). Subsequently cells were incubated at 37°C with a continuous supplement of 5% CO_2. COS-7 cells were grown until confluent and passaged at intervals of 3 days. At the point of confluency the culture supernatant of COS-7 cells was taken off and cells were washed in 10 ml PBS A (Ca^{2+} and Mg^{2+} free). 5 ml pre-warmed trypsin-ethylenediaminetetraacetic acid (EDTA; SIGMA) was added to the flask and incubated for 5 min at 37°C. To detach the cell layer the flask was tapped sharply for up to 5 times and 5 ml of complete DMEM were added to stop trypsin activity. Cells were transferred into a 50 ml Falcon tube, pelleted at 140 x *g* for 5 min and supernatant was discarded. Passaged cells were resuspended in 10 ml PBS A and 50 µl were mixed with 100 µl PBS A and 100 µl Trypan blue (Invitrogen™). 15 µl of the mix was now added to a haemocytometer and cells were counted according to standard protocols to determine the quantity of cells per ml. To continue culturing, cells were counted and $2x10^6$ cells were seeded into a 175 cm^2 flask.

2.2 Molecular Biology Techniques

All cloning for recombinant protein expression in *Escherichia coli* was performed using pET-30 Xa/LIC and pET-41 Ek/LIC Cloning Kits (Novagen) following the manufacturer's instructions. Additionally the pSecTag_2A vector was used for COS-7 cell expression and the pHH1-ΔSERA4 vector was employed for cloning of parasite transfection constructs. Bac-to-Bac® Baculovirus Expression System with pFastBac1™ vector (Invitrogen™) was used according to the manufacturer's instructions for cloning of recombinant bacmid DNA for subsequent insect cell expression. All insect cell transfection constructs were a kind gift from Michael Shea.

2.2.1 DNA modifying enzymes

Restriction endonucleases were obtained from New England Biolabs (NEB), Promega and Roche and all digests were performed using the manufacturer's instructions. DNA fragments were gel extracted with QIAquick® Gel Extraction Kit

(Qiagen). Antarctic phosphatase (NEB) was used to remove 5′ phosphate groups from the plasmid backbone.

2.2.2 Primers

Primers were diluted to a 100 µM stock in dH_20 and stored at -20ºC. Primers were used at a concentration of either 5 µM or 10 µM depending upon the enzyme used. A list of primers, the primer sequences and their use is shown in Table 5.

2.2.3 Polymerase chain reaction (PCR)

Platinum®Taq High fidelity DNA polymerase and Thermoprime *Taq* DNA polymerase (Thermo Scientific) were used for DNA amplification and screening respectively according to the manufacturer's instructions (Mullis, 1990). Oligonucleotides were obtained from Eurogentec and SIGMA. A synthetic SERA6 wild-type gene (synthSERA6 WT) recodonised for expression in *Pichia pastoris* (Geneart) was used for recombinant SERA6 expression. A synthetic PbSERA3 wild-type gene (synthPbSERA3 WT) recodonised for insect cell expression (Geneart) was used for recombinant PbSERA3 expression.

2.2.4 *E. coli* strains and transformation

The following *E. coli* strains were used for plasmid amplification, grown in lysogeny broth (LB): DH5α™ cells sub-cloning efficiency (Invitrogen™) or NovaBlue GigaSingles™ Competent Cells (Novagen). *E. coli* BL21-Gold DE3 cells (Stratagene), Shuffle®T7 Express *lysY* competent cells (NEB), Shuffle®T7 Express *lysY* competent cells containing the pG-KJE8 plasmid (generated during this work) and Origami™2-DE3 *pLysS* (Novagen) were used for protein expression (see Table 7). MAX Efficiency® DH10Bac™ competent *E. coli* were used for transposing the gene of interest from pFastBac1™ into the recombinant bacmid DNA (Invitrogen™) for insect cell expression.

2.2.5 Plasmid DNA preparation

All DNA was purified using QIAquick® Mini-prep Kit or QIAquick® Plasmid Maxi Kit (Qiagen). DNA concentration was measured either semi quantitatively by staining with SYBR®Safe gel stain (Invitrogen™) following electrophoresis on 0.7% agarose gels or in a Nanodrop spectrophotometer (Thermo Scientific). All steps were performed according to the manufacturer's instructions.

2.2.6 DNA ligation

DNA ligation was performed using either the Rapid DNA Ligation Kit (Roche) following the manufacturer's instructions or using pET-30 Xa/LIC Cloning Kit (Novagen) and pET-41 Ek/LIC Cloning Kit (Novagen).

2.2.7 Nucleotide sequencing

Nucleotide sequencing was performed by Beckman Coulter Genomics.

2.2.8 Integration PCR

To assess whether constructs had integrated into the targeted *P. falciparum* endogenous locus PCR was performed (see 2.2.3) using extracted genomic DNA from transfected parasite lines (see 2.1.3).

2.2.9 Southern hybridisation

2.2.9.1 DNA preparation

Genomic DNA extracted from transfected parasite lines (see 2.1.3) was digested with restriction enzymes to provide suitable sized DNA fragments for analysis. Digested DNA was separated on a 0.7% agarose gel (BioRad laboratories) containing a 1:20,000 dilution of SYBR® Safe DNA gel stain (Invitrogen™). DNA was exposed to UV on a transilluminator (UVP – Bio Doc-It) to nick the DNA in order to increase transfer efficiency. The agarose gel was incubated at room temperature with gentle agitation for 1 h in denaturing buffer (500 mM NaOH, 750 mM NaCl), rinsed in double-distilled water (ddH_2O) and incubated for 1 h in neutralising buffer (500 mM Tris-HCl pH 7.4, 750 mM NaCl) followed by a 20 min incubation in saline-sodium citrate buffer (SCC) containing 150 mM NaCl and 15 mM sodium citrate pH 7. The DNA was transferred onto a Hybond N+ membrane (Amersham Biosciences) overnight by capillary action transfer as described by (Sambrook J., 1989).

2.2.9.2 Hybridisation

The DNA probe was amplified from synthSERA6 by PCR using specific primers (see Table 5) and purified using QIAQuick® PCR purification kit (Qiagen), then labelled with anti-Digoxigenin-AP Conjugate (DIG) following the manufacturer's instructions using a DIG Luminescent Detection Kit (Roche). Hybridisation was done according to the manufacturer's instructions using a DIG Luminescent Detection Kit (Roche). The Southern blot was visualised by exposure to BioMaxTM MR film (Kodak®) at -80°C.

2.3 Vector construction

All vectors designed and used in this work are displayed in Table 6.

2.3.1 Constructs for recombinant protein expression in *E. coli*

(cloning and recombinant expression of PbS3C1 was conducted by Katarina Milutinovic)

Various plasmids for heterologous protein expression were designed based on SUB1 cleavage site predictions for SERA6. Two PfSUB1 processing sites and 1 Protease X processing site in SERA6 were previously predicted, based on known SERA5 processing sites (Yeoh *et al.*, 2007). Inserts were amplified by PCR from synthSERA6. Site directed mutagenesis (SDM) to replace the putative catalytic Cys644 with Ala644 in the synthSERA6 (referred to as synthSERA6 C644A) was conducted by using a QuickchangeII SDM kit (Agilent Technologies). Different constructs were obtained by amplification following the manufacturer's instructions and performed using pET-30 Xa/LIC Cloning Kit (Novagen) or pET-41 Ek/LIC Cloning Kit (see Table 6 and Figure 16). The PbSERA3 central 1 (rPbS3C1, between PbSERA3st1 and PbSERA3st2) was amplified either from synthPbSERA3 WT or from a modified version of the gene in which the putative catalytic Cys639 was replaced via SDM with an Ala residue (synthPbSERA3 C639A). DNA was amplified using primer pair PbS3C1_fwd_pET30 and PbS3C1_rev_pET30 and cloned using a pET-30 Xa/LIC Cloning Kit (Novagen; see Figure 17A/C and Table 5).

2.3.2 Constructs for recombinant transient expression of SERA6 and SERA6 C644A in COS-7 cells

DNA was amplified from synthSERA6 and synthSERA6 C644A. PCR was conducted according to standard procedure using the primer pair S6_FL_fwd_pSecTag2a_Sfi1 and S6_FL_rev_pSecTag2a_EcoR1, introducing a 5′-Sfi1 restriction site and a 3′-EcoR1 restriction site into the PCR product. The pSecTag2A backbone was pre-digested with Sfi1 followed by EcoR1 and the insert was ligated into the vector using standard ligation methods (see 2.2.6 and Figure 18).

2.3.3 Constructs for recombinant transient expression of PbSERA3 WT and PbSERA3 C639A mutant in *Tricoplusiu ni* BTI-Tn-5Bl-4 (Tn5) cells

(all constructs were a kind gift from Michael Shea and Fiona Hackett)

DNA for recombinant expression of full-length rPbSERA3 WT and full-length rPbSERA3 C639A in insect cells was amplified from synthPbSERA3 WT or synthPbSERA3 C639A respectively (see 2.3.1) using primers specific for full-length PbSERA3 and amplicons were subcloned into the pFastBac™ Vector (see Table 5 and Figure 17B). Recombinant baculovirus was prepared by recombination between the donor plasmid (pFastBac™) and the viral DNA (bacmid) via site-directed transposition

according to manufacturer's instruction for the Bac-to-Bac® Baculovirus Expression System (Invitrogen™). Virus stocks were generated by transfecting Sf9 cells (*Spodoptera frugiperda*, obtained from Invitrogen™) with bacmid DNA for PbSERA3 or PbSERA3 C639A and assayed using standard methodology.

2.3.4 Design of integration construct pHH-SERA6chim and derivatives

(all constructs were designed in collaboration with Michael Shea based on pET-30 Xa/LIC_S6FL)

Plasmids for integration into the parasite genomic *sera6* locus by single homologous recombination were based on the pHH1-ΔSERA4 vector (Child *et al.*, 2010). A 1020 bp region of the *sera6* gene downstream of the translational start site was amplified using primers SERA6-int-for and SERA6-int-rev (see Table 5) and ligated into the pHH1 backbone in place of the *sera4* coding sequence. The remainder of the *sera6* coding region (2406 bp) was excised from the synthSERA6 backbone and ligated in frame with the 3′ end of the *sera6* sequence in the pHH1 vector to generate a plasmid containing a region of genomic *sera6* to promote homologous recombination fused in frame to the rest of *sera6* in recodonised form to allow expression of the gene following integration. The transfection vectors were designed to integrate into the *sera6* locus to produce a chimeric gene that would either reconstitute a coding sequence encoding the wild-type SERA6 protein (construct pHH-SERA6chim), or that would express a series of mutant forms of SERA6 modified at the predicted cleavage sites SERA6st1 (construct pHH-SERA6chim_mutS1), SERA6st2 deletion (pHH-SERA6chimΔS2), SERA6st2 substitution (pHH-SERA6chim_mutS2), SERA6stX (pHH-SERA6chim_mutSx), or combinations of these mutations (see Table 6 and Figure 19) Mutations were introduced by SDM using the QuickchangeII SDM kit (Agilent Technologies) according to the manufacturer's protocol. The mutation for SERA6st1 used primer SERA6-17 and consisted of the replacement of residues Ala378-Gln379 with Leu368-Leu379. The mutation for SERA6st2 used primer SERA6-19 and consisted of the replacement of residues Gly935-Gln936 with Leu935-Leu936. The mutation for SERA6stX used primer SERA6-34 and consisted of the replacement of residues Tyr897-Phe898-Leu899 with Ala897-Ala898-Ala899.

2.4 *P. falciparum* transfections

Ring stage parasites at 5–10% parasitaemia were transfected by electroporation with 70 μg of plasmid DNA as described previously (Harris *et al.*, 2005). Selection for lines carrying the plasmid was performed with 10 nM WR99210 (Jacobus Pharmaceuticals, New Jersey, USA). Selection of integrants was promoted by cycling

on and off drug selection (2.5 nM WR99210). The parasites were cultured off drug for three weeks followed by drug selection.

2.5 Immunochemical Methods

2.5.1 Protein separation via Sodium Dodecyl Sulfate Polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970)

For disruption of protein disulfide bonds (reducing conditions), samples were diluted with 2x SDS sample buffer (62 mM Tris-HCl pH 6.8, 5% (v/v) glycerol, 3% (w/v) SDS, 100 mM dithiothreitol (DTT) and denatured at 95ºC for 5 min. Therefore all native protein interactions were completely eliminated, proteins denatured and finally charged with SDS but without entire protein dissolution (maintenance of primary structure). The proteins were separated via 7.5%, 10%, 12.5% or 15% resolving mini gels (Table 8), chromatographed at a constant current of 25 mA.

2.5.2 Immunoblotting (Towbin *et al.*, 1979;Sambrook J., 1989)

Qualitative immunodetection was obtained by the standard application of immunoblotting under wet conditions. Molecules were transferred to Hybond-C nitrocellulose (GE Healthcare) membrane at 150 mA overnight. The membrane was blocked with 5% milk-PBS, followed by three washing steps with PBS-Tween-20 (PBST). Primary antibodies were diluted in PBST, 1% (w/v) BSA and the blots were incubated with shaking for 90 min at RT or 4ºC overnight. Blots were then washed and incubated with PBST diluted secondary antibody which was coupled to horseradish peroxidise (HRP), for 30 min at room temperature. Enhanced chemiluminescence (ECL; Pierce; Ogata *et al.*, 1983) was used to visualise the targeted protein on a BioMaxTM MR film (Kodak®). Exposure times varied depending on the abundance of the protein under investigation and the dilution and specificity of the different antibodies (see Table 9).

2.5.3 Fixing and staining SDS-PAGE gels

To visualise proteins of interest SDS-PAGE gels were fixed and stained with 0.15% (w/v) Coomassie Brilliant Blue R-250, 45% (v/v) methanol, 45% (v/v) ddH_20, 9.85% (v/v) acetic acid for 5 min and destained with 5% (w/v) methanol and 10% (w/v) acetic acid overnight. After destaining was complete some SDS-PAGE gels were subjected to silver stain procedure with a PageSilver™ Silver Staining Kit (Fermentas) following manufacturer's instructions.

2.5.4 Antigen preparation for antibody generation against recombinant SERA6 central 1 (rS6C1) and recombinant C-terminus (rS6-Cterm)

Antibodies against the rS6C1 fragment and the recombinant C-terminus of SERA6 (see Table 9) were generated in house. Solubilised inclusion bodies (IB, see 2.6.4) were applied to a PD-10 Desalting Column equilibrated in PBS (Amersham Biosciences). The eluate was then injected into mice for polyclonal antibody generation (Malcolm Strath). Mouse sera were adsorbed against *E. coli* acetone powder to remove antibodies against bacterial contaminants, using 40 mg/ml acetone powder (Harlow *et al.*, 1988). Rabbit antibodies were generated using rS6C1 (see Figure 16B) from recombinant expression with pET-30 Xa/LIC_S6C1 (Harlan laboratories).

2.5.5 Sample preparation for N-terminal sequencing (Edman sequence analysis) of PfSUB1 processing sites of SERA6

(in collaboration with Michael Shea)

Recombinant SERA6 (rS6-FL; pET-30 Xa/LIC_S6FL), rS6-FLmutS1 (pET-30 Xa/LIC_S6FL_S1), rS6-FLmutS2 (pET-30 Xa/LIC_S6FL_S2), and rS6-FLmutS1S2 (pET-30 Xa/LIC_S6FL_S1S2) were expressed with N-terminal S- and 6xHis-tags using the pET-30 Xa/LIC system in *E. coli*. In all cases proteins were produced as IB which were washed and solubilised directly into SDS sample buffer in the presence of the reducing agent DTT (100 mM final concentration). Solubilised IB were separated by SDS-PAGE. The most abundant protein, corresponding to rS6-FL, was excised from the gel to separate it from breakdown products, then subjected to in-gel digestion with or without rPfSUB1 (Withers-Martinez *et al.*, 2002). Gel pieces were then boiled in SDS sample buffer to solubilise the protein fragments, and again subjected to SDS-PAGE, followed by transfer of the polypeptide fragments to polyvinylidene fluoride (PVDF) membrane. N-terminal sequencing was performed by the Protein and Nucleic Acid (PNAC) facility at the Department of Biochemistry, University of Cambridge.

2.5.6 Sample preparation for N-terminal sequencing (Edman sequence analysis) of PfSUB1 processing sites of rPbS3C1

The recombinant PbSERA3 central 1 domain (between PbSERA3site1 and site2 referred to as rPbS3C1) was incubated with rPfSUB1 (Withers-Martinez *et al.*, 2002) at 37°C for 2 h, concentrated 20 fold with a Vivaspin 500, 10 kDa cut-off concentrator (VWR), resuspended in 5x SDS sample buffer (containing 100 mM DTT) and subjected to SDS-PAGE. This step was subsequently followed by transfer to PVDF membrane. N-terminal sequencing was performed by PNAC.

2.5.7 Indirect immunofluorescence analysis (IFA)

Thin films of *P. falciparum* (various life stages) were air dried, fixed in 4% paraformaldehyde in PBS for 30 min at RT and permeabilised in 0.1% Triton X-100 (TX-100; SIGMA) for 10 min. Fixed slides were then washed three times with PBS for 10 min and blocked overnight at 4ºC in 3% (w/v) bovine serum albumin (BSA) in PBS. The following day slides were probed with the appropriate primary antibody (see Table 9) for 30 min at 37ºC and then washed three times for 10 min in PBS. Slides were stained with 4',6-diamidino-2-phenylindole (DAPI;GIBCO®), mounted in Citifluor, and images collected using AxioVision 3.1 software on an Axioplan 2 Imaging system (Zeiss) using a Plan-APOCHROMAT 1006/1.4 oil immersion objective (Harris *et al.*, 2005).

2.5.8 Immunoprecipitation

Highly synchronous *P. falciparum* schizonts were enriched with Percoll, pelleted and stored in aliquots at -80°C. Pellets were thawed into ice cold 10 mM Tris-HCl pH 8.2 containing cOmplete EDTA-free protease inhibitor cocktail (Roche) and supernatant was collected at 12,400 x *g* for 20 min at 4°C. Monoclonal antibody 1C9 (α-S6-Nterm), previously generated against the SERA6 N-terminus, was added to the parasite supernatant and incubated on ice for 1 h, manually mixing every 15 min. The sample was then added to 20 µl of packed Protein G Sepharose (GH Healthcare) shaking for 1 hr at 4°C and the suspension was then washed 3 times in PBS. Elutions were conducted as follows: 3 washes in 50 mM Tris-HCl, 5 mM EDTA, 0.5% (v/v) TX-100 pH 8.2, followed by 3 washes in 500 mM NaCl, 50 mM Tris-HCl, 5 mM EDTA, 0.5% (v/v) TX-100, 1 mg/ml bovine serum albumin pH 8.2 and 3 washes in 1 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, 0.5% (v/v) TX-100, 1 mg/ml BSA pH 8.2. Beads were finally solubilised into SDS sample buffer. All the fractions were subjected to SDS-PAGE.

2.6 Recombinant Protein Expression

2.6.1 Expression of recombinant protein in *E. coli*

Shuffle T7 Express *lysY* cells and BL21-Gold DE3 cells were transformed with several constructs in order to produce recombinant SERA6 protein (see Figure 16). A single picked colony was grown at 30°C overnight in 5 ml LB medium (containing the appropriate antibiotic) and then diluted 50-fold the following day. Shuffle T7 Express *lysY* cells were induced when cells had grown to an optical density of 0.6 with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16ºC overnight and BL21-Gold DE3 cells were induced with 0.5 mM IPTG at 18°C. Following induction, bacteria were pelleted at 2,000 x *g* for 20 min and lysed in an appropriate volume of BugBuster

Master Mix (Novagen). Table 7 shows all recombinant *E. coli* expression conditions used. Recombinant PbS3C1 was produced by transformation of BL21-Gold DE3 cells based on the rPbS3C1 construct designed in this study (see Table 6 and Figure 17C). Recombinant PbS3C1 was a kind gift of Katarina Milutinovic.

2.6.2 Transient expression of recombinant full-length SERA6 in COS-7 cells

The day before the transfection experiment cells were passaged and seeded at 1.2 x 10^7 cells per 175cm^2 flask in order that the cells were at least 80% confluency the following day. On the day of transfection, 42 μg of pSecTag_2A_S6FL or pSecTag_2A_S6FL_C/A plasmid DNA (see Table 6 and Figure 18) were incubated with 84 μl FuGENE® HD Transfection Reagent (Roche) and the volume was adjusted to 2.1 ml with DMEM medium (free of serum and Pen/Strep). The mix was incubated at RT for 15 min. In the meantime 40 ml of complete DMEM was added to each flask containing the attached cells. The DNA transfection reagent mix was added dropwise to the culture flasks and incubated for 7 h at 37°C in an atmosphere of 5% CO_2. After the incubation period the cells were washed 3 times with PBS A and finally 40 ml of serum free DMEM containing Pen/Strep, L-Glutamine, and 250 ng/μl tunicamycin (SIGMA) was added to each flask. After 4 days the supernatant was collected, residual cells were pelleted at 140 x *g* for 5 min and the supernatant was filtered through a 0.22 μm diameter filter (Corning). The supernatant was then concentrated 16 fold and stored in aliquots at -80°C until further use. The adherent cells were collected by using cell scrapers (Costar), pelleted at 140 x *g* for 5 min, and stored at -80°C until further use.

2.6.3 Transient expression of recombinant PbSERA3 in *Tricoplusiu ni* BTI-Tn-5Bl-4 (Tn 5) cells

(conducted by Fiona Hackett)

High FiveTM cells (Invitrogen™) were grown in ESF 921 medium and transfected with the under see 2.3.3 (also see Figure 17) produced viral stocks as described by (Withers-Martinez *et al.*, 2002). Tunicamycin was added at a final concentration of 500 ng/ml). Accumulation of secreted protein was monitored by Western blot using the mouse α-PbS3C1 antibody, raised against the central domain of PbSERA3 (kind gift from Volker Heussler). The medium was collected 72 h post-infection, cells were pelleted by centrifugation at 2,100 x *g* for 10 min, supplemented with glycerol to 10% (v/v) and stored at -80°C until further use.

2.6.4 Inclusion body solubilisation

All three recombinant SERA6 polypeptides rS6-FL, rS6C1 and rS6C2 were expressed in *E. coli* in a mainly insoluble form in IB. Following disruption of the bacteria with BugBuster Master Mix as described above, IB were washed twice in 50 mM Tris-HCl pH 8.1, 2% (v/v) TX-100, 500 mM NaCl, 5 mM DTT and solubilised in either 8 M urea or 6 M guanidine-HCl in a buffer containing 250 mM NaCl (500 mM respectively), 1 mM imidazole, 8 β-mercaptoethanol and 20 mM Tris-HCl pH 8.1. IB were solubilised at RT for 2 h and urea soluble and insoluble fractions were separated at 6,000 x *g* for 25 min.

2.6.5 Protein refolding

After solubilisation of IB in 8 M urea the denatured proteins had to be refolded in order to gain their native fold. Rapid dilution experiments (20, 50 or 100 fold dilutions) were conducted under a variety of different refolding conditions (see Table 10) using the Quick Refolding Kit (Pierce) according to the manufacturer´s instructions as well as refolding protocols based on the folding of SERA5 and *P. falciparum* falcipains described elsewhere (Sijwali *et al.*, 2001a;Hodder *et al.*, 2003). A range of redox conditions were used at 4°C, RT or 37°C, and with incubation periods of up to 5 days. Refolding status was assessed via clear native PAGE (CN-PAGE). Lower gel and upper gel buffers were prepared using 1.5 M Tris-HCl pH 8.8 and 500 mM Tris-HCl pH 6.8 respectively. 25 mM Tris-HCl pH 8.3 and 192 mM Glycine were used for the running buffer. Refolded samples were diluted into 2x native sample buffer, containing 125 mM Tris-HCl pH 6.8, 20% (w/v) glycerol and 0.01% Bromophenol blue. CN-PAGE gels were run for 2 h at a constant current of 25 mA. Proteins were then visualised with either Coomassie Brilliant Blue R-250 or the PageSilver™ Silver Staining Kit (Fermentas). Immunoblotting of CN-PAGE was conducted as described previously.

2.7 Protein purification

2.7.1 His-tag affinity purification of recombinant protein

Bacterial lysate (soluble BugBuster Master Mix fraction) or solubilised IB were incubated with Ni-NTA Superflow resin (Qiagen) rotating for 1 h at 4°C. The supernatant was removed after centrifugation at 12,400 x *g* for 1 min, and the resin was then washed 3 times by repeated resuspension in 300 mM NaCl, 50 mM NaH_2PO_4 and 20 mM imidazole pH 8. Proteins were eluted using either increasing concentrations of imidazole or acidic pH in buffers based on the washing buffer. Fragments from solubilised IB were eluted using either 8 M urea, 100 mM NaH_2PO4, 10 mM Tris-HCl at a pH of 5.9 for the first elution and pH 4.5 for the second elution. A different elution was

attempted using 8 M urea, 250 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 M imidazole, 20 mM β-mercaptoethanol.

2.7.2 Anion exchange purification of *E. coli* derived rS6-FL and rS6C1

Bacterial lysates (soluble BugBuster Master Mix fraction) were diluted 5 fold into either 20 mM HEPES pH 6.5 or 20 mM ethanolamine pH 9 and incubated on ice for 20 min with either SP-Sepharose® Fast Flow (Amersham Biosciences) or Q-Sepharose® (Amersham Biosciences) respectively. The suspensions were then centrifuged at 12,400 x *g* for 1 min and both supernatants were applied to Ni-NTA Superflow immediately, as described previously (see 2.7.1).

2.7.3 Purification of parasite-derived SERA6

To purify parasite derived full-length SERA6, 0.5 ml of frozen *P. falciparum* schizonts were thawed into 4 ml of ice-cold 10 mM Tris-HCl, pH 8.2, 10 mM EDTA, in the absence of protease inhibitors. Supernatant and membranes were separated by centrifugation at 10,000 x *g* for 15 min at 4°C. The cleared lysate was then filtered through 0.45 μm filter and 4 ml of supernatant was applied to the injection loop of a HiLoad 26/60 Superdex 200 prep-grade column (GE Healthcare), equilibrated as described elsewhere (Yeoh *et al.*, 2007). Determination of elution fractions containing full-length SERA6 was conducted via Western blotting. After gel filtration all fractions containing SERA6 were pooled, diluted 3 fold into 20 mM Tris- HCl pH 8.2 and subjected to anion exchange purification as described elsewhere (Yeoh *et al.*, 2007). The final product was referred to as parasite-derived semi-purified SERA6 (see Table 6).

2.7.4 Purification of rPbSERA3

The in 10% (v/v) glycerol stored culture supernatant was thawed and 400 ml was mixed with 400 ml 50 mM Tris-HCl pH 8.2. A HiPrep16/10Q XL column (GE Healthcare) was equilibrated in 50 mM Tris-HCl pH 8.2. Supernatant was loaded onto the column at a flow rate of 4 ml/min followed by a wash step with 100 ml 50 mM Tris-HCl pH 8.2. Recombinant PbSERA3 was eluted in 400 ml of 50 mM Tris-HCl pH 8.2 with a gradient from 0-400 mM NaCl in 10 ml fractions at a flow rate of 2 ml/min. Analysis of rPbSERA3 fractions was conducted via Western blotting. A HiLoad 26/60 Superdex 200 prep-grade column (GE Healthcare) was equilibrated overnight in 25 mM HEPES, 150 mM NaCl and 12 mM $CaCl_2$ pH 7.4. After anion-exchange chromatography rPbSERA3 containing fractions were pooled and 140 fold concentrated with 70 ml Centricon®Plus-70, cut-off 30 kDa (Millipore™), a volume of maximum 500 μl was loaded onto the column and sample was collected as described

elsewhere (Yeoh *et al.*, 2007). Western blot analysis followed to determine rPbSERA3 containing fractions.

2.7.5 Purification of SERA6-central domain from parasite culture supernatant

A parasite derived SERA6 central domain (SERA6-central; see Table 6) was pre-purified from *P. falciparum* culture supernatant via anion-exchange as described elsewhere (Silmon de Monerri *et al.*, 2011). It was then further purified through gel filtration using a HiLoad 26/60 Superdex 200 prep-grade column (GE Healthcare) with 20 mM Tris-HCl pH 8.2, 150 mM NaCl, 1 mM DTT, 1 mM EDTA. 5 ml fractions were collected with settings as described previously (see 2.7.3). The SERA6-central domain was detected by Western blot analysis.

2.8 Protein Biochemistry

2.8.1 Saponin and Streptolysin O treatment

Highly synchronous *P. falciparum* schizonts enriched with Percoll were divided into aliquots of 4 x 10^8 cells. Streptolysin O (SLO) was activated for 15 min at 37°C with 10 mM DTT. Four hemolytic units (HU) of SLO were then added to one aliquot of enriched schizonts, incubated for 6 min at RT and membrane and soluble fractions were separated by centrifugation at 12,400 x *g* for 10 min at 4°C. For saponin lysis, saponin was prepared as a 0.15% (v/v) stock solution in PBS. Schizont pellets were re-suspended in 10 cell volumes of sterile saponin and incubated at RT for 6 min. Membrane and soluble fractions were separated by centrifugation at 12,400 x *g* for 10 min at 4°C. Different parasite fractions were diluted in sample buffer and subjected to SDS-PAGE and Western blotting and probed with various antibodies (see Table 9).

2.8.2 PfSUB1 processing assays of parasite-derived semi-purified SERA6

Full-length SERA6 was purified as described previously (Yeoh *et al.*, 2007), concentrated and resuspended in PfSUB1 activity buffer containing 25 mM HEPES, 12 mM $CaCl_2$ pH 7.4 (see 2.7.3). Semi-purified full-length SERA6 was divided into aliquots of 80 µl and either no rPfSUB1, 10 µl rPfSUB1, 2 µl recombinant PfSUB1 prodomain (rPD) or 10 µl rPfSUB1 plus 2 µl rPD were added immediately before incubating at 37°C (Withers-Martinez *et al.*, 2002;Jean *et al.*, 2003). Samples were taken after 30 min, 1 h and 2 h and assayed via Western blotting.

2.8.3 SERA6 processing assay *in vivo*egress inhibition assay

2.8.3.1 Inhibition assays with trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane (E64)

Parasites at early schizont (2-4 nuclei) stage were enriched using Percoll, resuspended in 10 ml RPMI-1640 Albumax cell culture medium either with or without the addition of the cysteine protease inhibitor E64 (10 μM), and further cultured at 37°C. Samples for IFA and Western blotting were taken after 3 h, 6 h, 9 h and 12 h.

2.8.3.2 Inhibition assays with 4-[2-(fluorphenyl)-5-(1-methylpiperidine-4-yl)-1H pyrrol-3-yl]pyridine (compound 1)

Early schizonts (2-4 nuclei) were enriched with Percoll and resuspended in 10 ml RPMI-1640 Albumax cell culture medium which contained either 10 μM E64 and 2.5 μM compound 1 (Dvorin *et al.*, 2010) or E64 alone, and further cultured at 37°C. Compound 1 was stored at -20°C as a 10 mM stock solution in DMSO. Prior to use it was diluted 1:20 in DMSO to give a final concentration of 500 μM, then further diluted to 5 μM in RPMI-1640 Albumax cell culture medium, for assays. Samples for IFA and Western blotting were taken after 3 h, 6 h, 9 h and 12 h.

2.8.4 Fluorogenic peptide assays

SERA6 protease activity was assayed using peptidyl 7-amido-4-methylcoumarin hydrochloride (AMC) fluorogenic substrates (available from the host lab, see Table 11) (Kanaoka, 1977). A total of 15 different AMC peptides were diluted with 25 mM HEPES, 150 mM NaCl, 1mM DTT pH 6.5 to a concentration of 10 μM in a volume of 50 μl and distributed in a 96 well plate. Crude bacterial lysates (soluble BugBuster Master Mix fraction) of rS6C1 wild-type and the rS6C1 C644A mutant were diluted 10 fold into the same buffer and 50 μl was added to each well leaving the AMC peptides at a concentration of 5 μM in 100 μl total volume. Fluorescence measurements were taken every 3 min for 10 min followed by further readings after 1 h, 2 h and 3 h incubation at 37°C, using excitation and emission wavelengths of 380 nm and 460 nm respectively. Additionally 3 internally quenched peptides (see Table 11, A-C) were tested at a final concentration of 20 μM using excitation and emission wavelengths of 320 nm and 420 nm respectively. Additionally the parasite derived and with rPfSUB1 processed semi-purified SERA6 (see 2.8.2) was tested for activity against all the fluorogenic peptides under the above mentioned conditions.

2.8.5 SERA6-central activity assays

SERA6-central was purified from *P. falciparum* culture supernatant as described under 2.7.5.

2.8.5.1 Fluorogenic peptide assays

A total of 18 fluorogenic peptides were screened as described under 2.8.4. AMC and internally quenched peptides were resuspended at 20 μM (10 μM final) and 50 μM (25 μM final) respectively in 20 mM Tris-HCl pH 8.2, 150 mM NaCl, 1 mM DTT, 1 mM EDTA and 1:2 diluted with the gel filtration purified SERA6-central (see 2.7.5). Increase in fluorescence was measured as described under 2.8.4.

2.8.5.2 SERA6-central autocatalytic processing assays

Parasite-derived SERA6-central (see 2.7.5) was concentrated 5 fold using a Vivaspin 500, 10 kDa cut-off concentrator (VWR) and incubated at 37°C overnight. Samples were separated on 12.5% SDS-PAGE and possible autocatalytic processing was assessed by Western blot analysis and Coomassie stain.

2.8.5.3 SERA6-central proteolytic activity against RBC ghosts

RBC ghosts were obtained by resuspending 100 μl of RBCs in 1 ml 5 mM ice-cold KH_2PO_4, incubation on ice for 5 min and centrifugation at 12,400 x g for 20 min at 4°C. Phases were carefully separated by discarding the supernatant, then the milky coloured ghost containing fraction was transferred to a new tube and washed several times in PBS A. The washed ghost fraction was finally resuspended in 25 mM HEPES, pH 6.4, 1 mM DTT or 20 mM Tris-HCl pH 8.2, 150 mM NaCl, 1 mM DTT and incubated at 37°C with SERA6-cntral. Samples were taken after 10 min, 30 min, 90 min and 3 h incubation, and analysed by SDS-PAGE followed by silver stain with the PageSilver™ Silver Staining Kit.

2.8.5.4 SERA6-central activity screen with gelatin zymograms

12.5% SDS-PAGE gels were prepared according to standard procedure (see 2.5.1). During preparation of the running gel a gelatin stock solution (10 mg/ml in distilled H_20) was added to produce a final concentration of 0.1% protein (1 mg/ml). To screen for SERA6-central activity purified protein was mixed with 2x SDS sample buffer (without DTT), incubated for 10 min at RT, and electrophoresed on SDS PAGE using 1 x Tris-Glycine SDS running buffer according to the standard running conditions (see 2.5.1). The SDS-polyacrylamide gel was then incubated in 1x zymogram renaturing buffer (2.5% (v/v) TX-100 in ddH_20) for 30 min at RT (the buffer was exchanged once after 15 min). The zymogram renaturing buffer was discarded and replaced with one of

the following 1 x zymogram developing buffers: 25 mM HEPES pH 6.4, 1 mM DTT; 100 mM NaAc pH 5.5, 0.05% NP-40, 1 mM DTT; 20 mM Tris-HCl pH 8.2, 150 mM DTT, 1 mM EDTA; 20 mM Tris-HCl pH 8.2, 150 mM NaCl, 5 mM $CaCl_2$, 1 mM DTT, 1 mM EDTA. The gels were equilibrated for 30 min at RT with gentle agitation then the buffer replaced with fresh developing buffer and incubated at 37°C overnight. The following day the gels were stained with Coomassie Brilliant Blue R-250 for 30 min, a concentration of 0.5% (w/v) was prepared instead of the usual concentration of 0.1% dye (to enhance sensitivity). Gels were destained with appropriate Coomassie destaining solution. Areas of protease activity appeared as clear bands against a dark blue background where the protease has digested the gelatin substrate which then diffused out of the gel.

2.8.6 Recombinant PbSERA3 activity assays

Recombinant PbSERA3 WT and rPbSERA3 C639A were purified as described previously (see 2.7.4). 30 μl of purified protein was incubated for 2 h at 37°C with rPfSUB1 (Withers-Martinez *et al.*, 2002). Additionally several reagents or protease inhibitors were added individually, as follows: PD, 10, 20 or 100 μM E64, 20 μg/ml leupeptin, 20 μg/ml antipain, 1 μM PbICP (Rennenberg *et al.*, 2010) or 2 μM chagasin. After incubation with rPfSUB1, rPbSERA3 WT was incubated with rPfSUB1-processed rPbSERA3 C639A in the presence or absence of E64. Next rPfSUB1 processed rSERA3 WT was incubated with rPfSUB1 processed rPbS3C1 to determine whether putative active rPbSERA3 WT processes in *trans*. All samples were analysed by Western blotting.

Table 3: All in this work used tissues

Product	***Manufacturer***
O^+ Human RBCs	National Blood Bank
COS-7 cells	Invitrogen, UK

Table 4: Software and external services used in this work

Software	***Manufacturer***
Adobe Photoshop Elements 5.0	Adobe Systems Incorporated
Adobe Illustrator CS4	Adobe Systems Incorporated
DNAStar Lasergene 8	DNASTAR Inc., Madison, USA
AxioVision 3.1 software	Carl Zeiss Ltd., Hertfordshire
Primers	Eurogentec Ltd., Southampton
Sequencing	Beckman Coulter Genomics, Essex
N-terminal sequencing (Edman sequence analysis)	PNAC facility, Cambridge
Recodonised *sera6* and *Pbsera3* genes	Geneart AG, Germany

Table 5: List of all primers used in this work

Primer name	***Sequence (5'-3')***	***Function and Features***
SERA6-FL_fwd	ggtattgagggtcgcgagggtaacaaggttacagttatc	Forward amplification of full-length SERA6 for pET-30 Xa/LIC
SERA6-FL_rev	agaggagagttagagccttaaacataacagaagttacaatcttcacc	Reverse amplification of full-length SERA6 for pET-30 Xa/LIC
SERA6-P56_fwd	ggtattgagggtcgcgatgacttcaacccaaacgagtacaag	Forward amplification of SERA6 central 1/2(S6C1/2) for pET-30 Xa/LIC
SERA6-P56_rev	agaggagagttagagccttattgaccgtgaacgaaagcgttgtc	Reverse amplification of SERA6 central 1 (S6C1) for pET-30 Xa/LIC
SERA6-P50_rev	agaggagagttagagccttacaagaagtaaacgttcttcttccagga	Reverse amplification of SERA6 central 2 (S6C2) for pET-30 Xa/LIC
P56_N-term_rev	ctcgtcaacgttcttgatacagatagc	Reverse amplification of SERA6 central 1 (S6C1) N-terminus for pET-30 Xa/LIC
P56internal_fwd	gtgacactaactacgtttacgact	Primer for PCR screen and sequencing
S6_Cterm_fwd	ggtattgagggtcgctctgatgaatccgacgagactaac	Forward amplification of SERA6 C-terminus for pET-30 Xa/LIC
S6_FL_fwd_pET41_Ek_LIC	gacgacgacaagatggagggtaacaaggttacagttatc	Forward amplification of full-length SERA6 for pET-41 Ek/LIC
S6_FL_rev_pET41_Ek_LIC	gaggagaagcccggttaaacataacagaagttacaatcttcacc	Reverse amplification of full-length SERA6 for pET-41 Ek/LIC
S6C1/2_fwd_pET41_Ek_LIC	gacgacgacaagatggatgacttcaacccaaacgagtacaag	Forward amplification of SERA6 central 1/2 (S6C1/2) for pET-41 Ek/LIC
S6C1_rev_pET41_Ek_LIC	gaggagaagcccggttattgaccgtgaacgaaagcgttgtc	Reverse amplification of SERA6 central 1 (S6C1) for pET-41 Ek/LIC
S6_FL_fwd_pSecTag_2A_Sfi1	gagggcccagccggccgagggtaacaaggttacagttatct	Forward amplification of full-length SERA6 for COS-7 cell expression
S6_FL_rev_pSECTag_2 A_EcoR1	cccggattcttaaacataacagaagttacaatcttcac	Reverse amplification of full-length SERA6 for COS-7 cell expression
SERA6-9	gctcgagattatcgttatgtccattattatgtgatatc	Reverse amplification of SERA6 promoter
SERA6-15	gcagggaaactgtggtttggcttggatcttcgcttccaag	Forward amplification to mutate catalytic Cys644 to Ala644 in synthetic SERA6

Primer name	***Sequence (5′-3′)***	***Function and Features***
SERA6-16	cgtccctttgacaccaaaccgaacctagaag cgaaggttc	Reverse amplification to mutate catalytic Cys644 to Ala644 in synthetic SERA6
SERA6-17	gaagaaaaagatgaacgagatcaaggttaagttattggat gacttcaacccaaacgagtacaagttg	Forward amplification to mutate SERA6st1 from AQ to LL in synthetic SERA6
SERA6-18	cttctttttctacttgctctagttccaattcaataacctactgaag ttgggtttgctcatgttcaac	Reverse amplification to mutate SERA6st1 in synthetic SERA6
SERA6-19	gaacgattacgacaacgctttcgttcacttattatctgatgaa tccgacgagactaacaaa	Forward amplification to mutate SERA6st2 from GQ to LL in synthetic SERA6
SERA6-20	cttgctaatgctgttgcgaaagcaagtgaataatagactact taggctgctctgattgttt	Reverse amplification to mutate SERA6st2 in synthetic SERA6
SERA6_int_for	cagtatacatgatatgtccttaatttttagcttt	Forward amplification to integrate SERA6synth
SERA6_int_rev	ctcacccttgtgctcaacatcaatccttg	Reverse amplification to integrate SERA6synth
SERA6_5UTRb	aaaagtaaaagaccaaatgata	Forward amplification covering start codon endogenous
SERA6-34	gtcctggaagaagaacgttgccgccgcgagacacaaca ctgacttcatg	Forward amplification to mutate Protease X site from YFL to AAA in synthetic SERA6
SERA6-35	catgaagtcagtgttgtgtctcgcggcggcaacgttcttcttc caggac	Reverse amplification to mutate Protease X site in synthetic SERA6
SERA6-37	aagtaggagtcggacttagaa	Reverse amplification from 3′ to check for integration of SERA6st1
SERA6-38	ctgtgatatcacataataatggac	Forward amplification for Southern probe to check for integration of pHH-SERA6chim
SERA6-39	cagatgatgattcagttggaag	Reverse amplification for Southern probe to check for integration of pHH-SERA6chim
PbS3C1_fwd_pET30	Ggtattgagggtcgctccgacgaggactccaaggaaaa cgac	Forward amplification of PbSERA3 central1 (PbS3C1) for pET-30 Xa/LIC
PbS3C1_rev_pET30	agaggagagttagagccttactgtcctgacacgtccaggtt	Reverse amplification of PbSERA3 central1 for pET-30 Xa/LIC
PbSERA3-1	ggcaactgctccctggcctggctgttcgcttc	Forward amplification to mutate catalytic Cys639 to Ala639 in synthetic PbSERA3

Primer name	***Sequence (5'-3')***	***Function and Features***
PbSERA3-2	gaagcgaacagccaggccagggagcagttgcc	Reverse amplification to mutate catalytic Cys639 to Ala639 in synthetic PbSERA3
PbSERA3-3	gcgtctgctcaagaccaacg	Forward amplification to sequence over catalytic site in synthetic PbSERA3
PbSERA3-4	gacaagcttggtaccgcatgc	Reverse amplification including HindIII after stop in pFast-PbSERA3 for insect cell expression
PbSERA3-5	gtatacatttcttacatctatgccgacgaagtgatcgagtccc	Forward chimera of melittin and full-length PbSERA3
PbSERA3-6	gggactcgatcacttcgtcggcatagatgtaagaaatgtatac	Reverse chimera of melittin and full-length PbSERA3
PbSERA3-7	gttttccttgcccttctggtc	Reverse for sequencing over the start site in full-length PbSERA3

Table 6: Overview of all constructs used in this work

This table includes all constructs that were designed by the author, obtained as kind gifts and plasmids that were designed and cloned in collaboration with Michael Shea. The table is a complete overview of all constructs for recombinant protein expression and for single homologous recombination in the parasite and special features thereof. Vector and product name are different in some cases and product names are labelled in italics and bold. Recombinant proteins will referred to as their product names throughout this work.

Vector name/*Product name*	***Important features***	***Function***	***Source***
Protein expression constructs			
synthSERA6 WT	Synthetic and recodonised SERA6	Expression of recombinant full-length SERA6 in *P. pastoris*	This study and kind gift from Dr. Sharon Yeoh
synthSERA6 C644A	Synthetic and recodonised SERA6 C644A	Expression of recombinant full-length SERA6 catalytic Cys644→Ala644 mutant in *P. pastoris*	This study
pET-30 Xa/LIC_S6FL or ***rS6-FL***	N-terminal S- and His- tagged S6-FL T7 promotor	N-terminal S- and His-tagged full-length SERA6 for protein expression in *E. coli*	This study
pET-30 Xa/LIC_S6C1 or ***rS6C1***	N-terminal S- and His-tagged S6C1 T7 promotor	N-terminal S- and His-tagged SERA6 central domain 1 (S6C1) between SERA6st1 and SERA6st2 for protein expression in *E. coli* and antibody production	This study
pET-30 Xa/LIC_S6C1_C/A or ***rS6C1 C644A***	N-terminal S- and His-tagged S6C1-C644A T7 promotor	as rS6C1 with catalytic Cys644→Ala644 mutant	This study
pET-30 Xa/LIC_S6C2 or ***rS6C2***	N-terminal S- and His- tagged S6C2 T7 promotor	N-terminal S- and His- tagged SERA6 central domain 2 (S6C2) between SERA6st1 and SERA6stX	This study
pET-30 Xa/LIC_S6Cterm or ***rS6-Cterm***	N-terminal S- and His-tagged S6C-terminus T7 promotor	N-terminal S- and His-tagged SERA6 C-terminus between SERA6st2 and SERA6 very C-terminus for recombinant expression in *E. coli*	This study

Vector name/*Product name*	***Important features***	***Function***	***Source***
		and antibody production	
pET-41 Ek/LIC_S6FL ***not applied in any experiments***	N-terminal GST-, S- and His-tagged S6-FL T7 promotor	N-terminal GST-, S- and His-tagged full-length SERA6 for protein expression in *E. coli*	This study
pET-41 EK/LIC_S6C1 ***not applied in any experiments***	N-terminal GST-, S- and His-tagged S6C1 T7 promotor	N-terminal GST-, S- and His-tagged SERA6 central domain 1 (S6C1) between SERA6st1 and SERA6st2 for protein expression in *E. coli*	This study
pSecTag-2A_S6FL or ***rSERA6 WT***	N-terminal Ig κ-chainleader sequence SV40 promotor and origin CMV promoter	Full-length SERA6 wild-type for soluble protein expression in COS-7 cells	This study
pSecTag_2A_S6FL_C/A or ***rSERA6 C644A***	N-terminal Ig κ-chainleader sequence SV40 promotor and origin CMV promoter	Full-length SERA6 catalytic Cys644→Ala644 mutant for soluble protein expression in COS-7 cells	This study
pFastBac1™PbSERA3 or ***rPbSERA3 WT***	N-terminal PbSERA3 signal peptide	Full-length PbSERA3 for expression in Tn5 cells with its own signal peptide	Kind gift from Michael Shea and Fiona Hackett
pFastBac1™PbSERA3 C/A or ***rPbSERA3 C639A***	N-terminal PbSERA3 C639A signal peptide	Full-length PbSERA3 catalytic Cys639→Ala639 mutant for expression in Tn5 cells with its own signal peptide	Kind gift from Michael Shea
pET-30 Xa/LIC_S6FL_S1 or ***rS6-FLmutS1***	N-terminal S- and His-tagged S6-FL SERA6st1 mutant T7 promotor	Full-length SERA6st1 mutant for protein expression in *E. coli*	Collaboration with Michael Shea
pET-30 Xa/LIC_S6FL_S2 or ***rS6-FLmutS2***	N-terminal S- and His-tagged S6-FL SERA6st2 mutant T7 promotor	Full-length SERA6st2 mutant for protein expression in *E. coli*	Collaboration with Michael Shea
pET-30 Xa/LIC_S6FL_S1S2 or ***rS6-FLmutS1S2***	N-terminal S- and His-tagged S6-FL SERA6st1 and st2 mutant T7 promotor	Full-length SERA6st1 and st2 mutant for protein expression in *E. coli*	Collaboration with Michael Shea
Transfection constructs			

Vector name/*Product name*	***Important features***	***Function***	***Source***
pHH-SERA6chim	chimera between endogenous *sera6* and synthSERA6 encoding the C-terminal 2406 bp of SERA6	Single homologous recombination of SERA6 chimera: wild-type positive control	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chimC/A	chimera between endogenous sera6 and synthSERA6 C644A	Single homologous recombination of SERA6 chimera: catalytic Cys644→Ala644 mutant	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS1	chimera between endogenous *sera6* and synthSERA6 SERA6st1 mutant	Single homologous recombination of SERA6 chimera: SERA6st1 VKAQ DDFN P2/P1 replaced with LL	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim∆S2	chimera between endogenous *sera6* and synthSERA6 deletion of SERA6st2	Single homologous recombination of SERA6 chimera: SERA6st2 deleted P4-P1	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS1∆S2	chimera between endogenous *sera6* and synthSERA6 SERA6st1 mutant and deletion of SERA6st2	Single homologous recombination of SERA6 chimera: SERA6st1 deleted P2/P1 replaced with LL and SERA6st2 P4-P1 deleted	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutSx	chimera between endogenous *sera6* and synthSERA6 SERA6stX mutant	Single homologous recombination of SERA6 chimera: SERA6stX replaced P3-P1 with AAA	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS2	chimera between endogenous *sera6* and synthSERA6 SERA6st2 mutant	Single homologous recombination of SERA6 chimera: SERA6st2 P2/P1 replaced with LL	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)

Vector name/*Product name*	***Important features***	***Function***	***Source***
pHH-SERA6chim_mutS2Sx	chimera between endogenous *sera6* and synthSERA6 SERA6st2 and SERA6stX mutant	Single homologous recombination of SERA6 chimera: SERA6st2 P2/P1 replaced with LL and SERA6stX P3-P1 replaced with AAA	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS1S2	chimera between endogenous *sera6* and synthSERA6 SERA6st1 and SERA6st2 mutant	Single homologous recombination of SERA6 chimera: SERA6st1 and st2 P2/P1 replaced with LL	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS1S2Sx	chimera between endogenous *sera6* and synthSERA6 SERA6st1, SERA6st2 and SERA6stX mutant	Single homologous recombination of SERA6 chimera: SERA6st1 and SERA6st2 P2/P1 replaced with LL; SERA6stX P3-P1 replaced with AAA	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmid from (Child *et al.*, 2010)
pHH-SERA6chim_mutS1Sx	chimera between endogenous *sera6* and synthSERA6 SERA6st1 and SERA6stX mutant	Single homologous recombination of SERA6 chimera: SERA6st1 P2/P1 replaced with LL; SERA6stX P3-P1 replaced with AAA	Collaboration with Michael Shea based on pET-30 Xa/LIC constructs, pHH plasmids from (Child *et al.*, 2010)
Parasite-derived SERA6			
SERA6-central	Small central SERA6 fragment , contains papain-like domain	/	*P. falciparum* culture supernatant
Semi-purified SERA6	Full-length SERA6	/	*P. falciparum* late stage schizonts

Table 7: Overview of various *E. coli* expression conditions and respective plasmid constructs

***E. coli* strain**	***Constructs***	***Conditions***
BL21-Gold DE3	pET-30 Xa/LIC_S6FL pET-30 Xa/LIC_S6C1 pET-30 Xa/LIC_S6C2	0.1, 0.5, 1 mM IPTG 18°C, RT, 28°C, 37°C 4 or 6 hours
Origami™2-DE3 *pLysS*	pET-30 Xa/LIC_S6FL pET-30 Xa/LIC_S6C1 pET-30 Xa/LIC_S6C2	1 mM IPTG RT, 30°C, 37°C 4 or 6 hours
Shuffle®T7 Express *lysY*	pET-30 Xa/LIC_S6FL pET-30 Xa/LIC_S6C1 pET-30 Xa/LIC_S6C2 pET-41 Ek/LIC_S6FL pET-41 EK/LIC_S6C1	0.1, 0.5, 1 mM IPTG 16°C, 30°C, 37°C overnight In presence of a protease inhibitor cocktail (Roche) or any of the following protease inhibitor: Pepstatin, TPCK, Aprotinin, EDTA, E64, Leupeptin, PHMB, Chymostatin, Antipain or in presence of 20 mM proline or at pH 5.5
Shuffle®T7 Express *lysY* containing plasmid pG-KJE8 (dnaK-dnaJ-grpE-groES-groEL)	pET-30 Xa/LIC_S6FL pET-30 Xa/LIC_S6C1 pET-30 Xa/LIC_S6C1_C/A pET-30 Xa/LIC_S6Cterm pET-41 Ek/LIC_S6FL pET-41 EK/LIC_S6C1	0.1, 0.5, 1 mM IPTG 16°C, 30°C, 37°C overnight

TPCK: N-tosyl-L-phenylalanine chloromethyl ketone; PHMB: p-hydroxymercuribenzoate

Table 8: Recipes for SDS-PAGE stacking and resolving gels

Stacking gel	**4 %**
dH_2O	23.66 ml
Upper gel buffer 500 mM Tris-HCl pH 6.8	10 ml
Acrylamide solution (30% gel)	6.34 ml
10% APS (ammonium persulphate)	160 µl
TEMED (N,N,N',N'-Tetramethyl-1-,2-diaminomethane)	52 µl

Resolving gel	***7.5%***	***10%***	***12.5%***	***15%***
dH_2O	32 ml	26.6 ml	21.25 ml	16 ml
Lower gel buffer 1.5 M Tris-HCl pH 8.8	16 ml	16 ml	16 ml	16 ml
Acrylamide solution (30% gel)	16 ml	21.4 ml	26.75 ml	32 ml
10% APS (ammonium persulphate)	400 µl	400 µl	400 µl	400 µl
TEMED(N,N,N',N'-Tetramethyl-1-,2-diaminomethane)	60 µl	60 µl	60 µl	60 µl

Table 9: Overview of all antibodies produced and used throughout this work

Primary antibodies	***Type***	***Target***	***Dilution for IFA***	***Dilution for Western***	***Manufacturer***
Human X509	mAb	3D7 allele PfMSP1-42	1:10,000	N/A	In house
Mouse α-S6 central 1 (α-S6C1)	polyclonal	SERA6-S6 central 1 (S6C1)	1:1,000	1:500	In house
Rabbit α-S6 central 1 (α-S6C1)	polyclonal	SERA6-S6 central 1 (S6C1)	1:1,000	1:1,000	Harlan Laboratories
Mouse 1C9 (α-N-term)	mAb	SERA6-N terminus	N/A	1:1,000	In house
Mouse α-C-term	polyclonal	SERA6-C-terminus (after SERA6st2)	N/A	1:500	In house
Mouse 24C6.1F1	mAb	SERA5-P50	1:1,000	1:2,000	Jean-Francois Dubremetz
Rabbit α-PfSERA5	polyclonal	Full-length SERA5	1:1,000	1:10,000	Harlan Laboratories
Rabbit α-Aldolase	polyclonal	*Arabidopsis thaliana* fructose-1,6 bisphosphate aldolase	N/A	1:2,000	Agrisera
Rabbit α-AMA1	polyclonal	PfAMA1-Ectodomain	1:1,000	N/A	In house
Mouse α-PbSERA3 (α-PbS3C1)	polyclonal	PbSERA3 central aa620-874	N/A	1:3,000	Volker Heussler
Secondary antibodies					
Goat α-mouse-HRP conjugate	polyclonal	Mouse IgG	N/A	1:10,000	BioRad
Goat α-rabbit	polyclonal	Rabbit IgG	N/A	1:10,000	BioRad
Alexa Fluor 594 α-mouse	polyclonal	Mouse IgG	1:1,000	N/A	Invitrogen
Alexa Fluor 488 α-mouse	polyclonal	Mouse IgG	1:1,000	N/A	Invitrogen
Alexa Fluor 594 α-rabbit	Polyclonal	Rabbit IgG	1:1,000	N/A	Invitrogen
Alexa Fluor 488 α-rabbit	polyclonal	Rabbit IgG	1:1,000	N/A	Invitrogen
Alexa Fluor 594 α-human	polyclonal	Human IgG	1:1,000	N/A	Invitrogen

Table 10: Summary of all attempted rS6C1 refolding conditions

Buffer	***Reduced Glutathione (GSH)***	***Oxidised Glutathione (GSSG)***	***Nvoy***	***Polyethylene glycol (PEG)***	***Ethylenediaminetetraacetic acid (EDTA)***
Refolding Kit, Buffers 1-9	2 mM, 1mM	0.2 mM, 1 mM, 0.4 mM	/	/	with and without 1 mM
Refolding Kit, buffer 3	2 mM	0.2 mM	0.1%	0.2 mM, 0.13 mM, 0.7 mM	with and without 1 mM
(Hodder *et al.*, 2003)	1 mM (1:0.4)	0.25 mM(1:0.4)	0.1%	/	/
(Sijwali *et al.*, 2001a)	2 mM, 1 mM	1 mM	/	/	/

Table 11: List of all fluorogenic peptides used to screen for SERA6 activity

Peptide No.	***Peptide sequence***	***Peptide No.***	***Peptide sequence***
1	7-(Glutaryl-L-Phenylalanine)-AMC	11	Nα-benzoyl-Arg-AMC
2	Acetyl-Asp-Glu-Val-Asp-AMC	12	Ac-Asp-Glu-Val-Asp-AMC
3	Nα-benzoyl-DL-Arg-AMC	13	2-Gly-Gly-Leu-AMC
4	Nα-benzoyl-L-Arg-AMC	14	2-Gly-Met-Leu-AMC
5	N-t-BOC-Ile-Glu-Gly-Arg-AMC	15	Suc-Leu-Leu-Val-Tyr-AMC
6	Ac-Val-Glu-Ile-Asp-AMC	16	nGP-AMC
7	Z-Gly-Gly-Leu-AMC	A (TACE substrate IV)	Abz-LAQAVRSSSR-Dpa
8	BOC-Arg-Val-Arg-Arg-AMC	B (β-Secretease substrate VII)	Abz-VKMDAE-EDDnp
10	Ac-Trp-Glu-His-Asp-AMC	C (Cathepsin B substrate)	Abz-GIVRAK(Dnp)-OH

Figure 16: Schematic of all SERA6 constructs used for recombinant expression in *E. coli*

SynthSERA6 WT and C644A mutant genes with SERA6st1, SERA6st2 and SERA6stX are displayed with relevant amino acid positions of predicted protease processing sites (A). According to predicted processing sites various SERA6 fragments were amplified from synthSERA6 and expressed with pET-30 Xa/LIC and pET-41 EK/LIC. All in *E. coli* expressed proteins are named according to name of the protein product including various processing site mutants (B, C). The recombinant proteins were used for antibody production, protease activity screens and N-terminal sequencing.

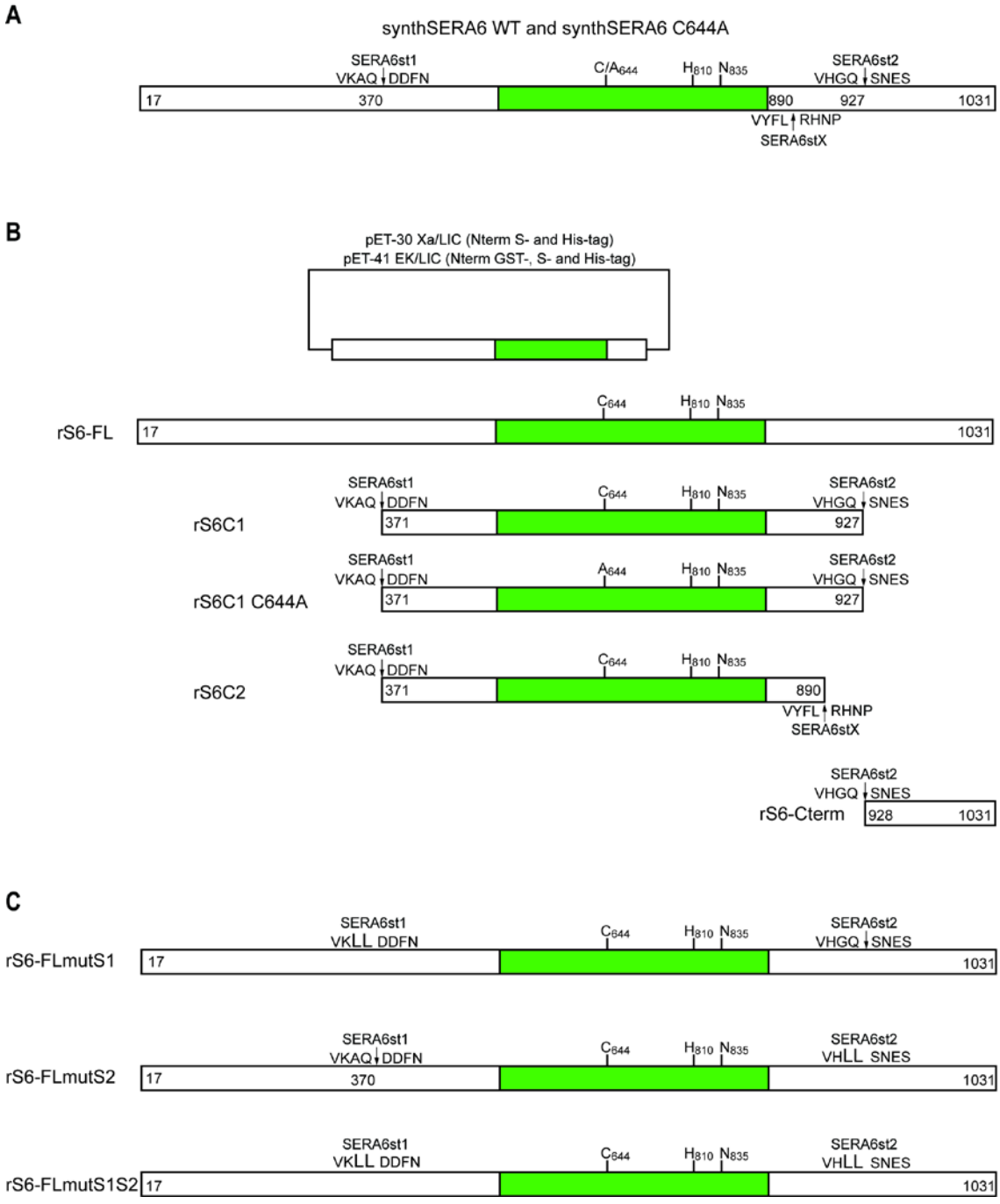

A
synthSERA6 WT and synthSERA6 C644A
SERA6st1
VKAQ↓DDFN
C/A644
H810 N835
SERA6st2
VHGQ↓SNES
17
370
890
927
1031
VYFL↑RHNP
SERA6stX
B
pET-30 Xa/LIC (Nterm S- and His-tag)
pET-41 EK/LIC (Nterm GST-, S- and His-tag)
rS6-FL
C644
H810 N835
17
1031
rS6C1
SERA6st1
VKAQ↓DDFN
371
927
rS6C1 C644A
A644
rS6C2
890
VYFL↑RHNP
SERA6stX
rS6-Cterm
928
C
rS6-FLmutS1
VKLL DDFN
rS6-FLmutS2
VHLL SNES
rS6-FLmutS1S2

Figure 17: Schematic of PbSERA3 constructs used for recombinant expression in *E. coli* and insect cells

SynthPbSERA3 WT and C639A mutant genes are displayed with relevant amino acid positions of predicted PbSERA3st1 and PbSERA3st2 (A). Full-length PbSERA3 was amplified from the synthPbSERA3 and expressed in Tn5 cells (B). PbS3C1 (fragment between PbSERA3st1 and PbSERA3st2) was expressed with pET-30 Xa/LIC (C). Schematics are labelled according to the name of the protein product. Recombinant material was then used for antibody production or protease activity screens.

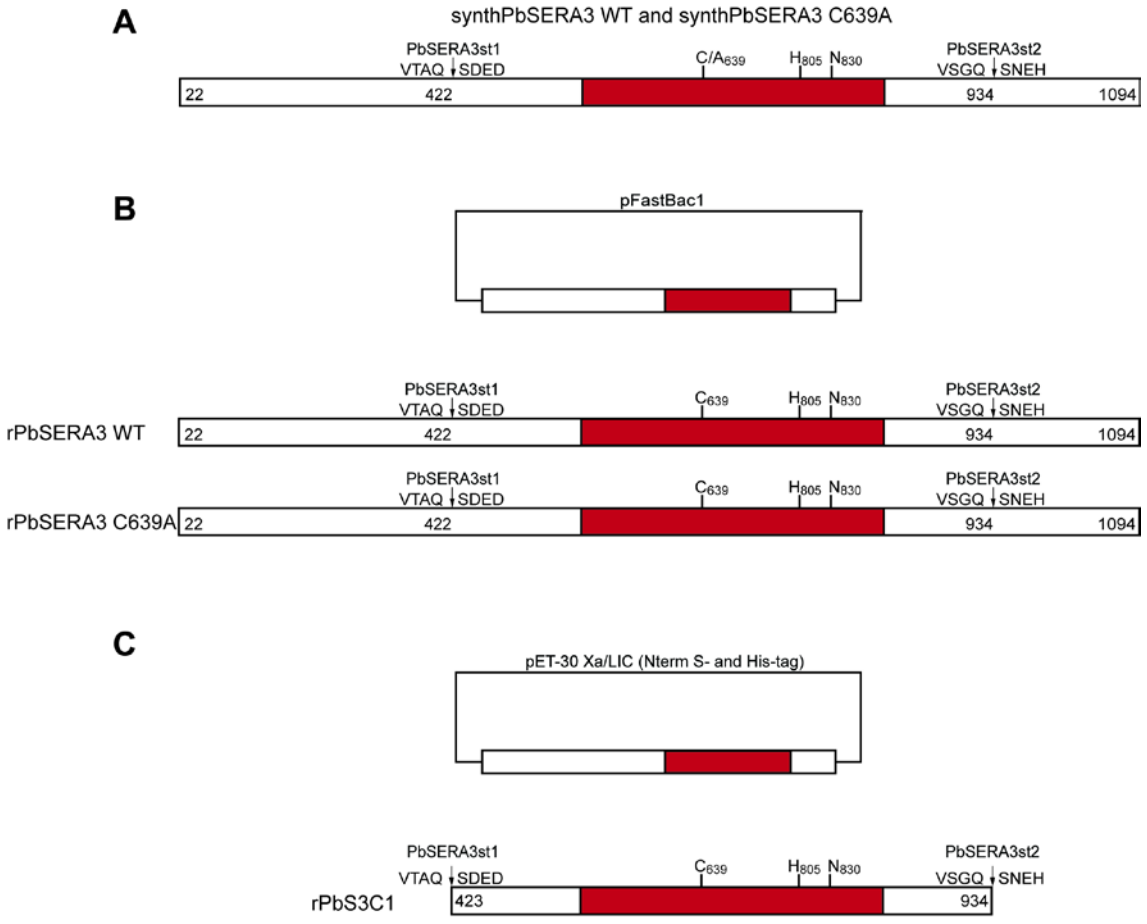

A
synthPbSERA3 WT and synthPbSERA3 C639A
PbSERA3st1
VTAQ ↓ SDED
C/A639
H805 N830
PbSERA3st2
VSGQ ↓ SNEH
22
422
934
1094
B
pFastBac1
rPbSERA3 WT
C639
rPbSERA3 C639A
C
pET-30 Xa/LIC (Nterm S- and His-tag)
rPbS3C1
423

Figure 18: Schematic of SERA6 expression construct in COS-7 cells

Full-length SERA6 was amplified from synthSERA6 WT and synthSERA6 C644A and expressed in COS-7 cells with pSecTag_2A. The recombinant protein was then used for PfSUB1 processing assays.

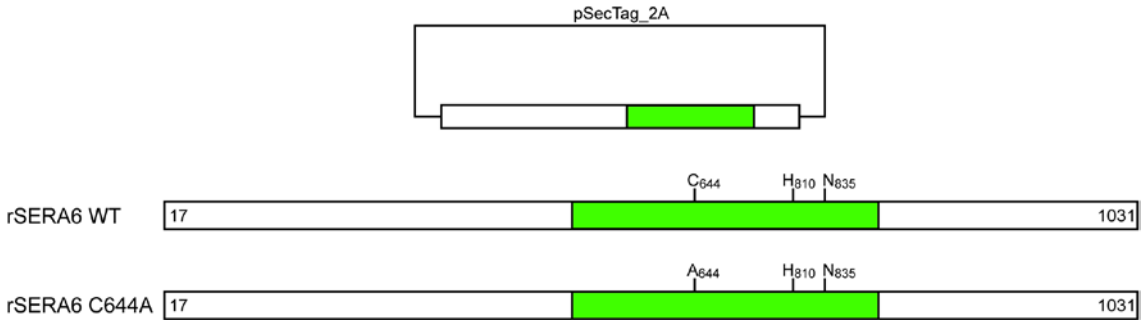
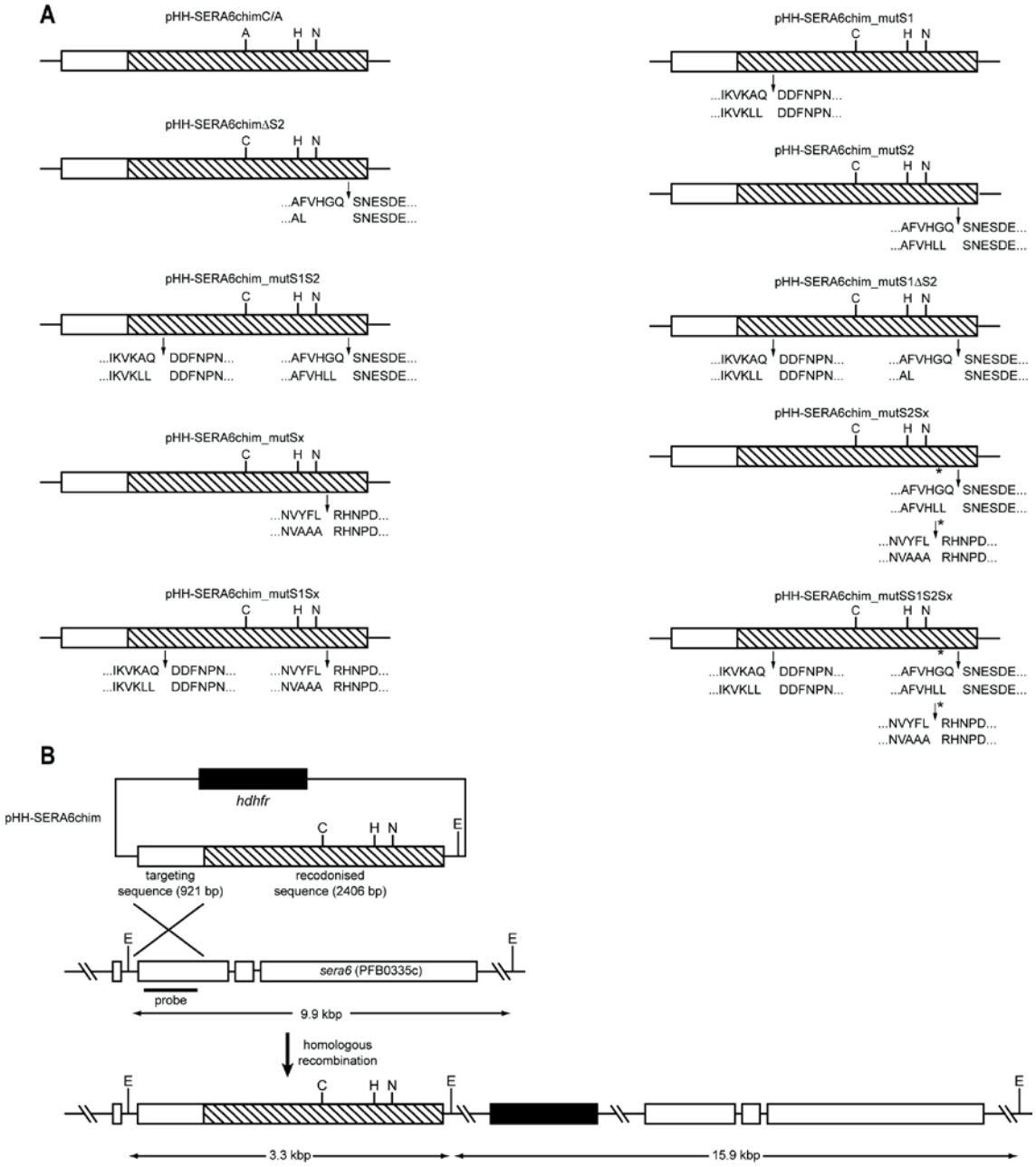

pSecTag_2A
C644
H810
N835
rSERA6 WT
17
1031
A644
H810
N835
rSERA6 C644A
17
1031

Figure 19: Schematic of all SERA6 transfection constructs for *P. falciparum* transfections

The various SERA6 integration constructs (designed by Michael Shea) are shown. To address the question of the importance of SERA6 processing, putative PfSUB1 and Protease X processing sites were modified in a manner predicted to block processing, either as a single-site mutants or in various combinations (A). A SERA6 wild-type chimera integrant served as control for each transfection. The integration strategy is based on single homologous recombination between the constructs and the endogenous *sera6* locus, producing a chimeric gene (B).


A
pHH-SERA6chimC/A
A H N
pHH-SERA6chimΔS2
C H N
...AFVHGQ SNESDE...
...AL SNESDE...
pHH-SERA6chim_mutS1S2
C H N
...IKVKAQ DDFNPN... ...AFVHGQ SNESDE...
...IKVKLL DDFNPN... ...AFVHLL SNESDE...
pHH-SERA6chim_mutSx
C H N
...NVYFL RHNPD...
...NVAAA RHNPD...
pHH-SERA6chim_mutS1Sx
C H N
...IKVKAQ DDFNPN... ...NVYFL RHNPD...
...IKVKLL DDFNPN... ...NVAAA RHNPD...
pHH-SERA6chim_mutS1
C H N
...IKVKAQ DDFNPN...
...IKVKLL DDFNPN...
pHH-SERA6chim_mutS2
C H N
...AFVHGQ SNESDE...
...AFVHLL SNESDE...
pHH-SERA6chim_mutS1ΔS2
C H N
...IKVKAQ DDFNPN... ...AFVHGQ SNESDE...
...IKVKLL DDFNPN... ...AL SNESDE...
pHH-SERA6chim_mutS2Sx
C H N
*
...AFVHGQ SNESDE...
...AFVHLL SNESDE...
*
...NVYFL RHNPD...
...NVAAA RHNPD...
pHH-SERA6chim_mutSS1S2Sx
C H N
*
...IKVKAQ DDFNPN... ...AFVHGQ SNESDE...
...IKVKLL DDFNPN... ...AFVHLL SNESDE...
*
...NVYFL RHNPD...
...NVAAA RHNPD...
B
pHH-SERA6chim
hdhfr
C H N E
targeting sequence (921 bp)
recodonised sequence (2406 bp)
E E
sera6 (PFB0335c)
probe
9.9 kbp
homologous recombination
E C H N E E
3.3 kbp
15.9 kbp

3 SERA6: a soluble PV protein

3.1 Introduction

As previously described, SERA5 is a soluble PV protein which undergoes extensive proteolytic processing by PfSUB1 just moments before egress (Knapp *et al.*, 1989;Aoki *et al.*, 2002;Miller *et al.*, 2002b;Yeoh *et al.*, 2007). It is the best studied and characterised *P. falciparum* SERA family member. However, very little information about the other SERA family members is available. Although SERA6 has a Cys residue in its active site and it appears to be essential for parasite survival (Miller *et al.*, 2002b), only a few studies have investigated its role during the parasite life-cycle. It is important to determine the subcellular localisation of SERA6 in order to gain more information about its function as it may be specific to its localisation. Although some previous studies have attempted to identify the exact localisation of SERA6, low-resolution pictures and cross-reactive antibodies complicated the matter (Knapp *et al.*, 1991;Aoki *et al.*, 2002;Yeoh *et al.*, 2007). SERA5 has successfully been expressed in bacterial cell systems for antibody production and crystallisation of the central domain (Hodder *et al.*, 2009). Therefore, it was attempted to express various fragments of SERA6 in bacterial cell systems in order to produce reliable and specific antibodies as a tool for investigating the role of SERA6. It was also considered that obtaining recombinantly expressed and correctly folded protein would be useful for activity studies.

Cysteine proteases often undergo extensive hydrolytic processing in order to become activated (Coulombe *et al.*, 1996). SERA5 is processed by PfSUB1 in at least two positions. Therefore, it was hypothesised for SERA5 that cleavage by PfSUB1 might be an activation event releasing a 56 kDa protease fragment. The last step is suggested to be followed by an additional processing step (mediated by Protease X) releasing a 50 kDa fragment indicating a possible inactivation event (Blackman, 2008). Alignments with SERA5 show that SERA6 might undergo a similar processing pattern which possibly indicates an activation event. The aim of this section of the study was to express various SERA6 fragments in a recombinant form in a heterologous system to then raise antibodies. These antibodies were then applied in biochemical and immunochemical approaches in order to establish that SERA6 localises to the PV.

3.2 Results

3.2.1 Bacterial expression of three recombinant SERA6 fragments (rS6-FL, rS6C1, rS6C2) results in insoluble proteins which were used for antibody production

The production of specific antibodies as a first step on the road to SERA6 characterisation was a primary aim of this work, for that SERA6 was expressed recombinantly. Based on the anticipated PfSUB1 cleavage sites (Yeoh *et al.*, 2007),

and one additional C-terminal cleavage site (unpublished data), three different SERA6 fragments rS6-FL, rS6C1, rS6C2 (representing the predicted processing products of SERA6) were aimed to be expressed, each with N-terminal 6xHis and S-tag for purification (see Figure 20).

An extensive variety of expression systems was approached in this study and in collaboration (Andrea Ruecker: *E. coli* expression systems, COS-7 cell expression system; Michael Shea: yeast expression system, insect cell expression systems, wheat germ cell-free system). All heterologous expression systems applied previously and in this study are displayed in Table 12. The pET-30 Xa/LIC was the system of choice for expression in *E. coli*. The advantage of this system lies within the ′3→5′ exonuclease activity of the T4 DNA polymerase which creates specific base pair overhangs that replace any ligation steps. In order to tackle the A/T-rich nature of the *P. falciparum* genome, the *sera6* gene was recodonised (for expression in *P. pastoris*) and all amplifications were conducted using the synthetic gene as a template. The insoluble fractions or IB usually harbour proteins that are misfolded or not folded at all. The expected molecular mass of the rS6-FL, rS6C1 and rS6C2 polypeptides are approximately 126 kDa, 70 kDa and 66 kDa respectively. All fragments were expected to display a slightly increased molecular weight due to the presence of N-terminal tags (all fragments possess one N-terminal 6xHis- and S-tag).

In a first attempt, *E. coli* BL21-Gold DE3 cells were transformed with constructs designed to express rS6-FL, rS6C1, rS6C2. Induction with IPTG for 4 hours at 37°C was approached in a first step to obtain soluble proteins of interest. The cells were collected and lysed with BugBuster Master Mix in order to separate soluble proteins from proteins expressed in IB (referred to as insoluble fraction) and subjected to SDS-PAGE followed by Coomassie staining; non-induced samples served as controls. Correctly refolded proteins should be detected in the soluble fraction. Although all three SERA6 fragments were expressed, the majority of each protein preparation appeared to be in inclusion bodies. A clear band appeared for each fragment in the insoluble fraction compared to the non-induced samples (see Figure 21A and B, fragments indicated with arrows a-c). To test whether any proteins, even small amounts, were expressed in the soluble fraction, the insoluble and soluble fractions were also analysed by Western blot, detecting with commercial α-His antibodies. The non-induced control showed that both rS6C1 and rS6C2 undergo leaky expression (see Figure 21C, asterisks), which can result in increased misfolding of recombinant proteins (Cabrita *et al.*, 2004). Signals for both rS6C1 and rS6C2 fragments appeared in both the insoluble and soluble fractions, suggesting that a small amount is soluble (see Figure 21D). Breakdown products were detected with the α-His antibodies in samples from both fragments and in insoluble fraction of rS6-FL. The recombinant

protein bands appear stronger in the insoluble fraction, indicating that the majority of recombinant protein is incorrectly folded. Only very weak Western blot signals were detected in the rS6-FL samples, which was surprising as a strong band of the expected molecular mass was detected on a Coomassie Blue stained SDS-polyacrylamide gel (see Figure 21B). This was possibly due to poor transfer of the rather large recombinant protein onto the nitrocellulose membrane.

In attempts to improve protein quantity and quality a range of induction conditions and times were examined. Additionally, protein expression was induced in the presence of protease inhibitors in order to prevent protease mediated breakdown (for an overview of all attempted conditions see Table 13). Shuffle T7 Express *lysY* cells express a chromosomal copy of a disulfide-isomerase in the bacterial cytoplasm which is known to improve folding of recombinant proteins and usually located in the periplasm. All conditions (as displayed in Table 13) were then analysed by a first screen on SDS-polyacrylamide gels stained with Coomassie Blue followed by Western blotting with a commercial α-His antibody. The results of this experiment showed that expression in Shuffle T7 Express *lysY* cells successfully improved the solubility of rS6C1 and rS6C2. Both SERA6 fragments showed increased expression levels in Shuffle T7 Express *lysY* cells compared to BL21-Gold DE3 cells, with a reduction of breakdown products upon overnight induction at 18°C (see Figure 22). Despite multiple attempts to obtain soluble protein using different expression conditions and different bacterial expression strains, the recombinant proteins were again predominantly recovered in IB. Therefore, it was decided to solubilise and refold the denatured proteins. Refolding of the purified protein was attempted using a variety of refolding conditions based on previous refolding efforts published for SERA5, falcipains and general refolding methods and will be analysed in chapter 5 (Sijwali *et al.*, 2001a;Hodder *et al.*, 2003). Despite numerous attempts using different solubilisation protocols, no correctly refolded SERA6 was recovered (see chapter 5). The focus for purification was therefore directed towards the small amount of protein which was expressed in soluble form in the bacterial cytosol (see Figure 22).

Regardless of the protein solubility, different protein purification techniques (Nickel-Histidine chelation, cation-anion exchange and gel filtration) failed to purify rSERA6-FL and rS6C1 from large scale expression preparations (see Figure 23A). The previous observation that the expression of rSERA6 was accompanied by breakdown products appeared to have a strong impact on the purification process, and the breakdown products were purified preferentially instead of the actual protein (see Figure 23A, arrow). It was therefore decided to solubilise rSERA6 from IB in 8 M urea to at least obtain enough material in order to raise antibodies (see Figure 23B). The rS6C1 was solubilised in 8 M urea, applied to a PD-10 desalting column and eluted

with PBS. Further purification from the eluted material failed as SERA6 was copurified with other proteins and breakdown products (data not shown). Therefore the PD-10 eluate and unpurified rS6C1 was used to immunise mice and rabbits (immunisation of mice was conducted by Malcolm Strath), thus raising antibodies against the central SERA6 domain, which contains the putative catalytic domain. As purification of rS6C1 antigen before immunisation was insufficient, the resulting mouse and rabbit polyclonal α-S6C1 antibodies were pre-absorbed with *E. coli* acetone powder in order to minimise any cross-reactivity with *E. coli* derived antigens.

3.2.2 Bacterial expression of the recombinant SERA6 C-terminus results in soluble protein and antibody production

The N- and C-terminal fragments of SERA5 produced upon PfSUB1-mediated processing are translocated to the merozoite surface where they are thought to play an important role in processes following egress (Li *et al.*, 2002b). The N- and C-termini of SERA6 might undergo a similar fate. Antibodies against the extreme N-terminus, the conserved C-terminus and the putative catalytic domain were therefore required for differential characterisation and analysis of SERA6 localisation and proteolytic processing. A construct was designed to express a small C-terminal 6xHis- and S-tagged peptide (rS6-Cterm) based on the putative PfSUB1 processing sites (see Figure 20). This was expressed in a soluble form and was successfully purified by Nickel-Histidine chelation followed by buffer exchange, before immunisation of mice (see Figure 24A, asterisk). Although the rS6-Cterm is only 18 kDa in molecular weight, a prominent breakdown product was observed. Western blot performed with a commercial α-His antibody revealed that N-terminal tags were present on the breakdown fragments (see Figure 24B). However, the breakdown products did not interfere with protein purification; they were copurified and coinjected into mice and antibodies were raised for further experiments (immunisation of mice was conducted by Malcolm Strath). The resulting polyclonal α-S6-Cterm antibodies displayed a very low titer constituting challenges for further experiments.

3.2.3 SERA6 antibodies recognise various SERA6 fragments

In the previous sections antibodies were raised against the recombinant SERA6 central region between the two predicted PfSUB1 cleavage sites (mouse and rabbit polyclonal α-S6C1), and the recombinant C-terminus (mouse polyclonal α-S6-Cterm) (see Figure 25A).

The aim of this section was to test and verify the functionality and specificity of the previously produced antibodies on recombinant SERA6 and parasite extracts. A polyclonal antibody raised against the N-terminal region of SERA6 was produced prior

to this work by immunisation with an *E. coli*-derived GST-fused to an N-terminal region of SERA6 and based on a construct as described elsewhere (Miller *et al.*, 2002b) (conducted by Michael Shea). A monoclonal SERA6 antibody 1C9 (α-S6-Nterm, serotyped IgG) was also available in the lab (work conducted by Malcolm Strath). In total, three different antibodies were generated to investigate the fate of SERA6. All are expected to recognise full-length SERA6 before the putative cleavage by PfSUB1, as well as their cognate fragments after PfSUB1 processing.

To examine reactivity of the antibodies with parasite-derived SERA6, total schizont lysates were separated by SDS-PAGE and probed in Western blots with the antibodies. As shown in Figure 25B, the three different antibodies all recognised a single band at approximately 135 kDa in the schizont lysates, presumably corresponding to full-length SERA6. The expected molecular weight of SERA6 is approximately 120 kDa, but these results agree with earlier findings that SERA6 migrates slower than expected on SDS-PAGE (Knapp *et al.*, 1991). In order to validate the specificity of the antibodies for SERA6, the nitrocellulose membranes were stripped and reprobed with a monoclonal antibody against SERA5 (mAb 24C6.1F1); SERA5 was detected as a band clearly distinct from SERA6 (see Figure 25B). Weak bands appearing in some of the schizont lysates are ascribed to low-level processing of SERA5 (see Figure 25B, asterisks). As SERA5 is the most abundant SERA, the results suggest that the antibodies specifically recognise SERA6 with no cross-reactivity with other members of the SERA family.

3.2.4 Antibodies against SERA6 are of low efficiency in immunoprecipitation

To analyse the processing by PfSUB1 and to determine SERA6 protease activity, an efficient system for the purification of SERA6 was required. The next set of experiments therefore focused on whether the α-SERA6 antibodies could be used to immunoprecipitate SERA6 from parasite lysates.

Frozen schizonts were subjected to hypotonic lysis in ice cold buffer and the soluble fraction was incubated with the monoclonal α-S6-Nterm antibody. To bind antibody-antigen complexes, the sample was then incubated with protein G beads and bound antigen was eluted in a buffer containing 500 mM NaCl. All samples were then analysed by Western blot with α-S6C1 antibodies. As shown in Figure 26A, although not all parasite-derived SERA6 was bound by the antibodies, a small amount of protein was eluted (see Figure 26A). Reprobing the same nitrocellulose membrane with polyclonal SERA5 antibodies (24C6.1F1) detected a second band below SERA6 (see Figure 26B). A signal for SERA6 was still visible in Western blots with SERA5 antibodies as both antibodies were raised in mice. As expected in a non cross-linked assay, the IgG of the N-terminal mouse antibody was also eluted and detected with

polyclonal mouse α-S6C1 (see Figure 26A, asterisk). At this point the rabbit polyclonal was not yet available.

In a next step it was attempted to cross-link the N-terminal antibody and the protein G sepharose. Cross-linking between primary amines of protein G and monoclonal α-S6-Nterm was unsuccessful as the antigen binding site of the antibody appeared to be inaccessible to SERA6 after cross-linking (data not shown). Therefore, to ensure that SERA6 would be eluted without the corresponding antibody, different elution conditions were explored showing that the antibody and SERA6 could be recovered in different elution fractions (see Figure 26C/D).

Despite successful separation of the antibody and SERA6, immunoprecipitation never recovered sufficient material in order to progress with the characterisation of SERA6 processing or activity assays. In conclusion it can be said that the immunoprecipitation was unsuccessful and other purification methods have to be applied in order to obtain purified SERA6 protein.

3.2.5 Immunofluorescence assays of late stage schizonts suggest that SERA6 is localised in the PV

Having produced antibodies specific for SERA6, the next part of the work set about using them to examine the subcellular localisation of the protein in the parasite. As a first approach, the polyclonal rabbit α-S6C1 antibodies were used to probe fixed smears of mature schizonts in immunofluorescence analysis (IFA). A marker for the merozoite surface protein MSP1 and antibodies against SERA5, a PV protein, were included in the IFA experiments. In a first set of experiments fixed schizont smears were dually labelled with SERA5 (mAb 24C6.1F1) and SERA6. The results of this set of experiments indicated that in highly segmented schizonts, the α-SERA6 signal largely shows a similar pattern as the α-SERA5 signal, possibly suggesting a PV localisation (see Figure 27A/B). However, signals for α-SERA5 and α-SERA6 were not completely identical; whilst the SERA6 signal was observed in a similar pattern to the α-SERA5 antibody, towards the central region of the schizonts, in 71% (of 106 schizonts) of all dual-labelled schizonts the α-SERA6 signal also showed a strong accumulation towards the edges of the PV, where the α-SERA5 signal was not visible. These results suggest a possible PV localisation of SERA6 but with a slightly different distribution from SERA5 in this compartment.

Next antibodies against SERA6 and MSP1 were used for dual labelling of fixed schizont smears. Antibodies against MSP1 label the surface of merozoites within the schizont and therefore permit a visualisation of schizont morphology similar to the differential interference contrast microscopy used by Glushakova *et al*. (Glushakova *et*

al., 2005;Glushakova *et al.*, 2010). Highly segmented but tightly packed schizonts may be indicative of an intact PVM, while highly segmented schizonts with more diffuse merozoite localisations might indicate that the PVM has already ruptured. SERA6 shows a broadly different pattern to MSP1 with only slight similarities of the antibody signals due perhaps to the nature of the iRBC morphology. In segmented but densely packed schizonts SERA6 is distributed throughout the PV (see Figure 27C). In IFAs performed on highly segmented schizonts with merozoites spread further apart, suggesting rupture of PVM but EPM still intact, SERA6 could not be detected in 21% (of 102 schizonts) of these schizonts respectively (see Figure 27D). The data therefore suggested that SERA6 may escape from the iRBC after PVM rupture but prior the rupture of the EPM. The observed pattern is due to the lack of SERA6 and not antibody effects as of two juxtaposed schizonts one showed a clear SERA6 signal, but the one which appeared to have egressed, did not. The latter showed MSP1 staining indicating that the antibodies were distributed efficiently (see Figure 27E). Therefore the question arose whether SERA6 underwent processing to full-fill its putative proteolytic function and then escaped from the iRBC before the rupture of the PVM. Colocalisation studies using antibodies against the N- and C-termini may elucidate this matter. However, it was not possible to obtain clear IFA signals with these antibodies presumably as both the N- and C-terminal antibody titers were very low. These antibodies only recognise SERA6 in IFAs of parasite strains expressing an additional episomal copy of SERA6 (conducted by Michael Shea).

Finally, the central domain of SERA6 is not detectable (using rabbit α-S6C1 antibodies) on the surface of free merozoites (see Figure 27F). To summarise: SERA6 appeared to be located to the PV in a manner similar to SERA5. However, a small percentage shows subcellular localisation distinct from SERA5, with distribution possibly towards the PVM. Antibodies against the SERA6 central domain did not show any signal in very late stage parasites, suggesting that the processed SERA6 was not retained in the PV after PVM breakdown.

3.2.5.1 Early results suggest that the subcellular distribution of SERA6 is conserved in *P. knowlesi*

SERA6 appears to be conserved throughout *Plasmodium* species and IFAs were performed on fixed schizont smears of a different *Plasmodium* species, *P. knowlesi*, using the rabbit α-S6C1 antibodies. *P. knowlesi* was successfully cultured and maintained in human RBCs (Robert Moon, unpublished) and highly synchronous late stage parasites were used for IFAs. Interestingly, the observed localisations were similar to *P. falciparum* SERA6, suggesting high levels of conservation between the central domain of SERA6 and its *P. knowlesi* orthologue (see Figure 28). Parasites

were dually labelled with rabbit α-S6C1 and mouse α-PkMSP19 (which labelled the surface of merozoites). Although the staining pattern of PkMSP19 was somewhat diffuse, some merozoite surface boundaries could be observed (see Figure 28). The initial results from IFAs on *P. knowlesi* schizonts suggest that the subcellular distribution of SERA6 may be conserved in other *Plasmodium* species.

3.2.6 Biochemical fractionation of infected RBCs confirms SERA6 as a soluble PV protein

Colocalisation studies of SERA5 and SERA6 through IFAs suggested that SERA6 is located in the PV albeit in a somewhat different pattern to SERA5, a soluble PV protein. In this section it was attempted to confirm these previous findings by a complementary approach. To do this two different chemical iRBC fractionation steps were performed and then analysed by Western blot. The approach was based on the biochemical properties of the pore-forming streptolysin O (SLO), which only introduces pores into the EPM and saponin, a cholesterol-dependent detergent which forms pores in the EPM and the PVM (for a detailed overview see Figure 29A). Depending on the subcellular localisation of SERA6 it can then be detected in different fractions

Highly synchronous *P. falciparum* schizont cultures were enriched by Percoll gradient centrifugation and divided into aliquots of $4x10^8$ cells. Either saponin or SLO were added for 6 min at RT. All soluble *P. falciparum* proteins resident in the PV should appear in the soluble fraction after saponin treatment but in the insoluble fraction after SLO treatment (see Figure 29A for detailed overview). The different fractions obtained after centrifugation were then separated by SDS-PAGE and examined by Western blot. The Western blots were probed with several different antibodies. SERA5 (mAb 24C6.1F1) was used as a control for a soluble PV protein and SERA6 (rabbit polyclonal α-S6C1) was expected to be detected in a similar manner. Antibodies raised against aldolase of *Arabidopsis thaliana* also recognise *P. falciparum* aldolase (a soluble parasite cytosol protein). The α-Aldolase antibody was therefore applied as control to ensure that the parasite plasma membrane (PPM) was still intact. Additionally to test that SLO did form pores in the EPM one SDS-polyacrylamide gel was stained with Coomassie Blue to determine the release of haemoglobin.

As seen in Figure 29B signals for full-length SERA6 and processed fragments were detected mainly in the insoluble SLO fraction and the soluble saponin fraction (see Figure 29B, top left panel). SERA5 display a similar profile to SERA6 (see Figure 29B, top right panel). Both proteins were detected in the fractions which contain soluble PV proteins. Aldolase could only be detected in the insoluble fractions, guaranteeing that no contents of the parasite cytosol could be found in any other fraction (see Figure 29B, bottom left panel). Additionally, the release of haemoglobin into the soluble SLO

and saponin fractions (see Figure 29B, bottom right) verified the completion of the respective permeabilisation steps. Therefore, the results agree with previous findings and suggest that SERA6 is mainly present as a soluble protein in the PV. However, a small proportion of SERA6 can consistently (in 6 independent experiments) be detected in the membrane fractions after saponin treatment, possibly suggesting association of SERA6 with membranes. This observation is supported by the result that SERA6 appeared to translocate towards the edges of possibly the PVM in IFAs.

Here, the earlier findings of Knapp *et al*. that SERA6 accumulates in the PV were confirmed, by IFA and biochemical fractionation of iRBCs (Knapp *et al.*, 1991). SERA6 showed PV localisation in late stage schizonts with a partial membrane association.

3.3 Discussion

In this part of the study antibodies specific to three different SERA6 fragments were designed, produced and verified as a tool for further characterisation. In addition, work in this part of the study aimed to produce soluble recombinant SERA6 for activity and processing studies. Finally and most importantly the subcellular localisation of SERA6 to the PV was confirmed by using two distinct approaches: IFAs and biochemical fractionation of iRBCs.

The data suggests that recombinant expression of SERA6 might be a long and winding road. It has previously been shown that expression of cysteine-rich malaria proteins and putative malarial proteases coincides very often with insolubility and expression or refolding challenges (Shenai *et al.*, 2000;Rosenthal, 2004;Kumar *et al.*, 2007). A small number of malarial proteases which appeared challenging for recombinant expression at the very first attempts have later been produced in large amounts. However, it cannot be assumed that this will be the case for the SERA family. After many years of work across several laboratories, SERA5 has only been expressed as its putative catalytic domain (Hodder *et al.*, 2003;Hodder *et al.*, 2009) and twice as full-length protein (Li *et al.*, 2002a; Robert Stallmach, unpublished). The high proportion of Cys residues in SERA6 is likely to have caused the insolubility. Previous work in the host lab included not only a diverse range of expression conditions in one system but investigation of a variety of distinct pro-and eukaryotic expression systems for SERA6, without success. A final approach was taken by expressing recombinant SERA6 in the green monkey kidney cell lines (COS-7) in order to obtain protein for activity assays (see chapter 5).

However, expression trials have not been completely futile. First and most importantly, antibodies against various SERA6 fragments were produced and these constituted the key tools which were required for further SERA6 characterisation. As it

was very challenging to purify recombinant full-length SERA6, the application of the monoclonal antibody for immunoprecipitation and purification of parasite SERA6 became even more important. It was possible to purify parasite SERA6 from highly synchronous late stage schizonts and to separate the protein from the antibody used for the immunoprecipitation. Unfortunately, the monoclonal N-terminal antibody could not be immobilised on Protein G or any other resins to facilitate purification as this presumably prevented epitope binding to SERA6. In addition, not enough material was recovered from these approaches. The focus therefore shifted towards the purification of parasite-derived SERA6 through large scale gel filtration followed by anion-exchange chromatography. These appeared more promising as parasite-derived SERA5 was successfully purified with these methods (Yeoh *et al.*, 2007) (see chapter 4).

It has to be emphasised that SERA6 was previously, but not very convincingly, shown to be localised to the PV as a homologue to SERA5 (Knapp *et al.*, 1991). To gain certainty, the localisation was repeated in this study with two distinct approaches. The results showed that SERA6 showed a similar subcellular but slightly distinct localisation to SERA5. Although both proteins displayed overlapping signal distribution in a PV lumen-like pattern, SERA6 additionally showed a concentration in its distribution towards what seemed to be the edges of the iRBC, perhaps suggesting a tendency to associate with the PVM. Interestingly, biochemical fractionation supported this, with a small amount of SERA6 being retained in the membrane fraction. SERA6 staining was observed in late stage schizonts which appeared to have an intact PVM. As described above, late stage schizonts with a possibly ruptured PVM but intact EPM lost the signal for the central SERA6 domain. These results suggest a putative spatial and temporal co-incidence between the putative SERA6 processing and the liberation of merozoites from iRBCs. This was further confirmed as no α-SERA6 signal was detected in IFAs of late stage schizonts which were egress-inhibited with E64 (data not shown). Although E64 was shown to inhibit parasite egress some data suggest that it does not prevent the breakdown of the PVM (Salmon *et al.*, 2001;Gelhaus *et al.*, 2005;Glushakova *et al.*, 2008).

The timing of the breakdown of the PVM during or prior egress is unknown. Many theories have been established which describe the fusion between the EPM and the PVM or even a breakdown of the EPM prior to PVM rupture preceding parasite egress (Winograd *et al.*, 1999;Salmon *et al.*, 2001;Soni *et al.*, 2005). More recently very convincing evidence indicated that the PVM disintegrates before EPM rupture (Glushakova *et al.*, 2005;Glushakova *et al.*, 2008). Most studies are based on labelling of PVM proteins as a marker for egress. However, in order to address and follow the question of the events leading to parasite release a *P. falciparum* endogenous PVM

marker has to be established. *P. falciparum* exported protein 1 (EXP1) which is inserted in the PVM could serve as an appropriate candidate (Simmons *et al.*, 1987;Kara *et al.*, 1988;Spielmann *et al.*, 2006). Parasite lines expressing for example an epitope-tagged EXP1 could allow for colocalisation studies with SERA6 in order to explore whether SERA6 is associated with the PVM. Additionally, the question of whether the central SERA6 signal disappears with PVM rupture could be further explored. The results suggest that, considering SERA6 as an active cysteine protease located in the PV and possibly escaping before EPM rupture, it may be involved in a cascade leading to PVM breakdown. However, an accumulation towards the edges could also imply an interaction of SERA6 with the EPM.

Intriguingly, the subcellular localisation of SERA6 and its *P. knowlesi* orthologue appeared to be conserved. The orthologue also shows a distribution across the PV with an accumulation towards the edges. As *P. knowlesi* merozoites are considerably larger than *P. falciparum* merozoites this could contribute towards the understanding of the subcellular localisation of a SERA6 orthologue (Lee *et al.*, 2009). Future work is required to analyse whether the observed pattern is representing PkSERA4, the SERA6 orthologue of *P. knowlesi*, or whether the α-S6C1 indeed cross-reacts with another *P. knowlesi* orthologue. SERA6 and PkSERA4 share 51% identity that is mainly found within the central domain, including the papain-like protease domain. This perhaps suggests similar functions for the central domains of SERA6 and PkSERA4, indicating once more that PkSERA4 might serve as a model to identify the protease activity of SERA6 and its orthologues. The analysis of the subcellular localisation of PkSERA4 however requires further verification through additional biochemical analysis. Localisation of another SERA6 orthologue, PbSERA3, also revealed the protein to be hosted within the space of the PV in hepatic schizonts (Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009) further supporting the findings from this chapter. Work in this section also aimed to address the fate of the SERA6 N- and C-termini. Despite antibody titers in both the N- and C-terminal antibodies being sufficient to recognise full-length SERA6 in parasite-derived material via Western blotting, they failed to detect the SERA6 in IFAs (data not shown). Using antibodies against the SERA6 central domain, SERA6 or at least the central domain were never detected on the surface of merozoites. Due to the limitations that arise by using optical microscopy in combination with the small diameter of schizont stage iRBCs, final conclusions on the colocalisation of SERA5 and SERA6 cannot be drawn. Future colocalisation studies for example with high resolution immuno electron microscopy will be required to finalise the analysis of the exact spatial distribution of SERA6.

It was therefore decided to examine the subcellular localisation of SERA6 with an additional approach. Infected RBCs were fractionated with saponin or SLO

(Ansorge *et al.*, 1996) to separate the parasite from contents of the PV and the host cell cytosol or from the host cell cytosol respectively. The data showed that SERA6 is indeed located to the PV and therefore confirmed the IFA results. Therefore it can be summarised that SERA6 is a soluble PV protein which is in agreement with previous findings (Knapp *et al.*, 1991;Aoki *et al.*, 2002). However, a small proportion of SERA6 remained associated with membranes in the biochemical fractionation of iRBCs, hinting at the above suggested membrane association.

The PV is a parasite compartment of particular interest. Its function is poorly understood but it is known that it serves as a transit and sorting compartment for protein transport and nutritional uptake (Ansorge *et al.*, 1996;Ansorge *et al.*, 1997;Lingelbach *et al.*, 1998). A recent study suggested that a large proportion of proteins located within the PV are proteases (Nyalwidhe *et al.*, 2006). Its surrounding membrane, the PVM, constitutes the contact interface between the parasite and the host cell. The protease PfSUB1, which plays a major role in egress, was the first protease found to be discharged into the PV (Yeoh *et al.*, 2007). The PV lumen is small in diameter, making it perfect for processes which are spatially and temporally tightly regulated such as egress (Eksi *et al.*, 2011). Recent studies demonstrated that the cysteine protease inhibitor E64 does not prevent breakdown of the PVM but inhibits egress and rupture of the EPM (Glushakova *et al.*, 2008) which theoretically excludes a role of cysteine proteases in this step. In this study, it was also observed that the cysteine protease inhibitor E64 may not prevent the rupture of the PVM which would suggest that SERA6 might not be involved in the disintegration of the PVM.

Homology modelling of the SERA5 papain-like domain with the protease domains of falcipain 2 and falcipain 3 (Chrislaine Withers-Martinez, unpublished) has suggested that the entrance to the S2 pocket in SERA5 is restricted by the aromatic ring of the Tyr735 residue. This Tyr residue is conserved in SERA6 (Tyr783) and PbSERA3 (Tyr780). E64 requires binding to the S subsites to then covalently bind to the active site Cys (Varughese *et al.*, 1989). The Tyr residue in the S2-pocket restricts the access of E64 to the catalytic Cys residue. E64 is unlikely to fit into the SERA6 active trough and is therefore an unsuitable inhibitor. Unusually high concentrations of E64 may therefore be required for inhibition of SERA6. Together with the findings from this section, it is possible that SERA6 is firstly expressed as a soluble PV protein and upon processing by PfSUB1 it may relocate to its place of function to trigger the breakdown of the PVM or EPM as an active cysteine protease.

Table 12: Overview of all heterologous expression systems used for SERA6

Expression system	***Expression construct***	***Application***
P. pastoris (Michael Shea, preceding this work)	rS6-FL: Cterm 6xHis-tag rS6C1: no tag rS6C2: no tag	Protein for antibody production and activity assays
E. coli	rS6-FL: Nterm 6xHis-tag, S-tag or GST-tag rS6C1: Nterm 6xHis-tag, S-tag or GST-tag rS6C1 C644A: Nterm 6xHis-tag, S-tag rS6C2 Nterm: 6xHis-tag, S-tag or GST-tag	Protein for antibody production and activity assays
Wheat germ cell-free system (Michael Shea)	rS6C2: Nterm 6xHis-tag, S-tag	Protein for activity assays
Insect cells (Michael Shea)	rS6-FL: Cterm 6xHis-tag and V5 epitope rS6C1: Cterm 6xHis-tag and V5 epitope rS6C2: Cterm 6xHis-tag and V5 epitope rS6-FL: no tag	Protein for antibody production and activity assays
COS-7 cells	rS6-FL: no tag rS6-FL C644A: no tag	Protein for activity assays

Table 13: Summary of all in this part applied bacterial expression systems and expression conditions

E. coli *strain*	***Construct***	***Expression conditions***
BL21-DE3 Gold	rS6-FL: Nterm 6xHis-tag, S-tag or GST-tag or Thioredoxin-tag rS6C1: Nterm 6xHis-tag, S-tag or GST-tag rS6C1 C644A: Nterm 6xHis-tag, S-tag rS6C2: Nterm 6xHis-tag, S-tag	0.1, 0.5, 1 mM IPTG 18°C, RT, 28°C, 37°C 4 or 6 hours
Origami2-DE3 *pLysS*	rS6-FL: Nterm 6xHis-tag, S-tag rS6C1: Nterm 6xHis-tag, S-tag rS6C2: Nterm 6xHis-tag, S-tag	1 mM IPTG RT, 30°C, 37°C 4 or 6 hours
Shuffle ®T7 Express *lysY*	rS6-FL: Nterm 6xHis-tag, S-tag rS6C1: Nterm 6xHis-tag, S-tag rS6C1 C644A: Nterm 6xHis-tag, S-tag rS6C2: Nterm 6xHis-tag, S-tag	0.1, 0.5, 1 mM IPTG 16°C, 30°C, 37°C overnight In presence of a protease inhibitor cocktail (Roche) or any of the following protease inhibitor: Pepstatin, TPCK, Aprotinin, EDTA, E64, Leupeptin, pHMB, Chymostatin, Antipain or in presence of 20 mM proline or at pH 5.5
Shuffle ® T7 Express *lysY* containing the plasmid pG-KJE8 (dnaK, dnaJ, grpE, groES, groEL)	rS6C1: Nterm 6xHis-tag, S-tag	0.1, 0.5, 1 mM IPTG 16°C, 30°C, 37°C overnight

TPCK: N-tosyl-L-phenylalanine chloromethyl ketone; pHMB: p-hydroxymercuribenzoate

Figure 20: All SERA6 constructs for generation of antibodies

The four different constructs used in this section were based on the predicted SERA6 PfSUB1 cleavage sites (SERA6st1 and SERA6st2) and a third cleavage by an unknown leupeptin sensitive protease called Protease X (SERA6stX). The putative papain-like catalytic domain is indicated with its three active site residues: Cysteine (Cys644), Histidine (His810) and Asparagine (Asn835). The N-terminal 6xHis- and S-tags for purification are shown in red. The molecular weights of the constructs are approximately 126 kDa (rS6-FL), 70 kDa (rS6C1), 66 kDa (rS6C2) and 18 kDa (rS6-Cterm). The fragments were amplified from a recodonised SERA6 gene between the indicated amino acid residues (aa). N: N-terminus; C: C-terminus

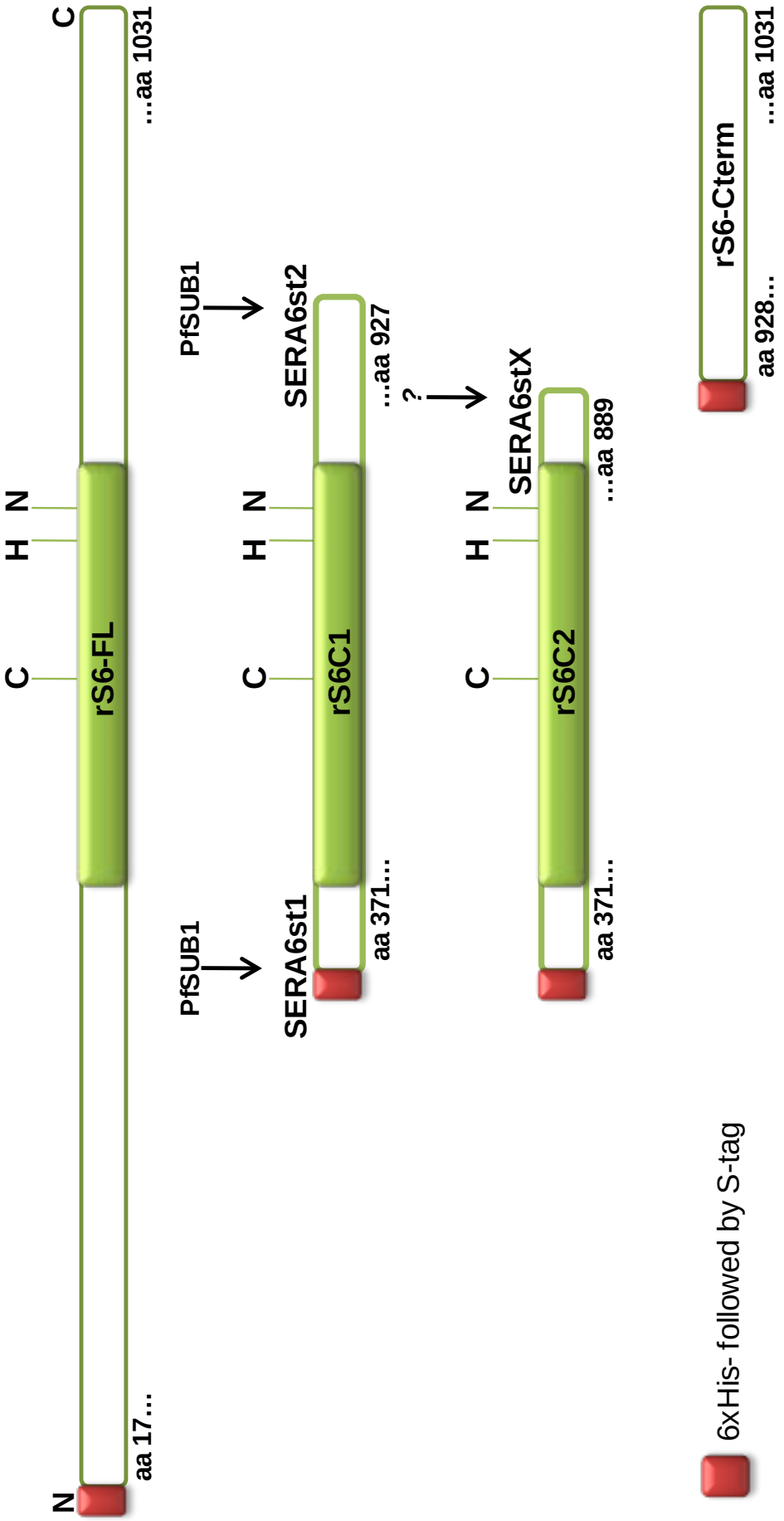

C
...aa 1031
N
H
C
rS6-FL
N
aa 17...
PfSUB1
SERA6st2
...aa 927
N
H
C
rS6C1
SERA6st1
aa 371...
?
SERA6stX
...aa 889
N
H
C
rS6C2
aa 371...
rS6-Cterm
...aa 1031
aa 928...
6xHis- followed by S-tag

Figure 21: Three different SERA6 fragments are expressed in the insoluble fraction in *E. coli*

BL21-Gold DE3 cells were transformed with expression plasmids for expression of rS6-FL, rS6C1 and rS6C2 and induced overnight. The cells were then lysed and soluble fraction (sol) and inclusion bodies (IB, insoluble fraction) were subjected to SDS-PAGE, and stained with Coomassie Blue to visualise expression patterns. Non-induced samples, taken prior to addition of IPTG, were separated via SDS-PAGE and stained with Coomassie Blue (A). All SERA6 fragments are mainly expressed in IB (B, a-c); breakdown products can be observed, particularly for rS6C1 and rS6C2. Western blots of the non-induced control showed that *E. coli* transformed with constructs for expression of rS6C1 and rS6C2 undergo leaky expression, showing that the fragments were expressed before the induction with IPTG (C, asterisks). This can cause increased misfolding of recombinant proteins. It appeared that rS6C1 and rS6C2 were expressed as insoluble and soluble proteins as bands for both fragments were present in both fractions (D). The breakdown fragments appeared to be carrying the N-terminal tags as they were detected with a commercial α-His antibody (D). Recombinant S6-FL could not be detected in insoluble or soluble form suggesting poor transfer from the SDS-polyacrylamide gel to the nitrocellulose membrane. IB: inclusion bodies; sol: soluble fraction after lysis with BugBuster Master Mix

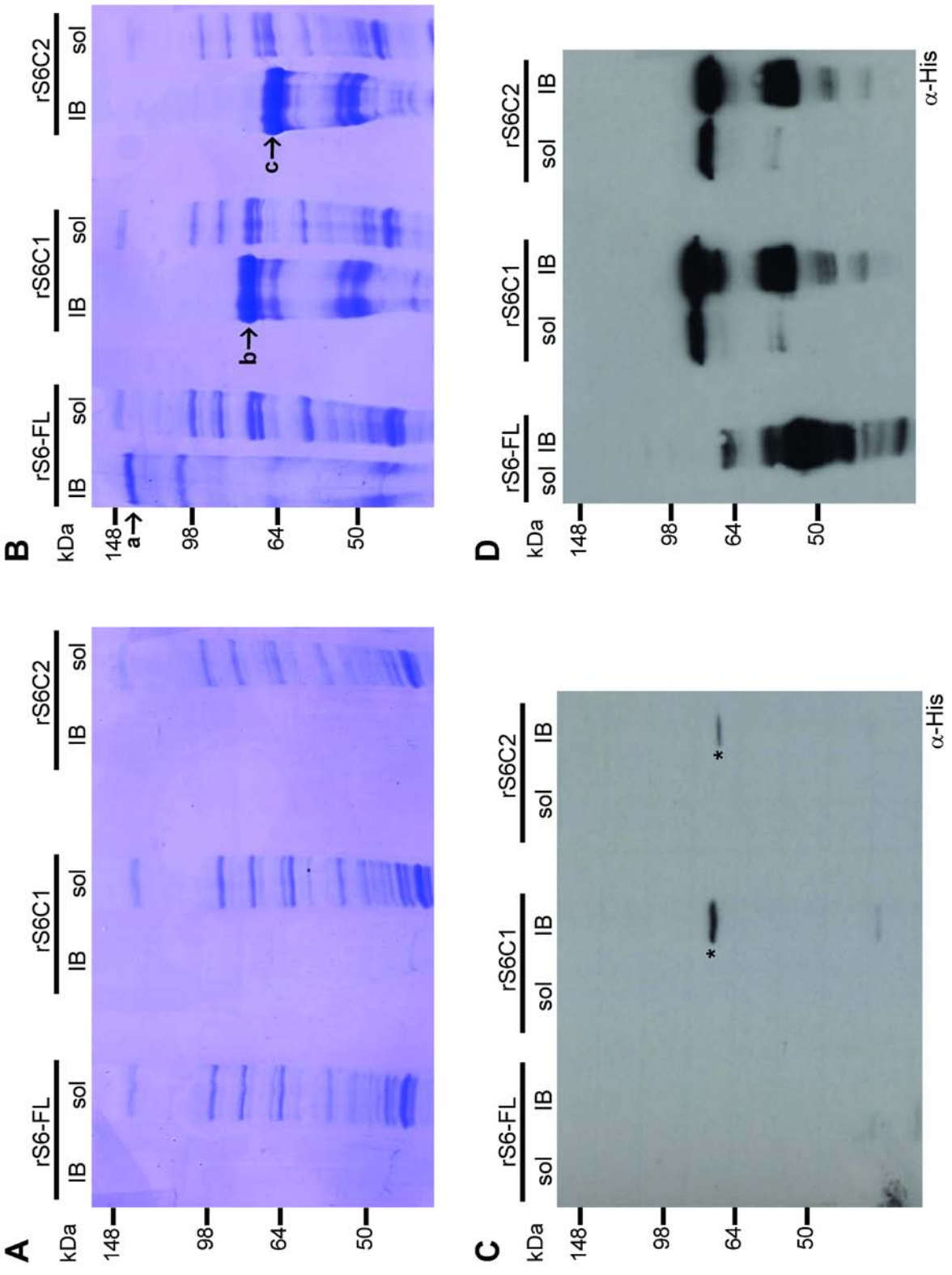

A
B
C
D
kDa
148
98
64
50
rS6-FL
rS6C1
rS6C2
IB
sol
a→
b→
c→
*
α-His

Figure 22: Expression of rS6C1 and rS6C2 in Shuffle cells improved protein solubility

In order to improve the solubility of the recombinant protein, Shuffle T7 Express *lysY* and BL21-Gold DE3 cells harbouring the relevant expression plasmids were induced under various conditions. The soluble fractions (sol) obtained after lysis with BugBuster Master Mix were subjected to SDS-PAGE (identical loadings) and detected by Western blotting with commercial α-His antibodies. sol: soluble fraction after BugBuster Master Mix lysis

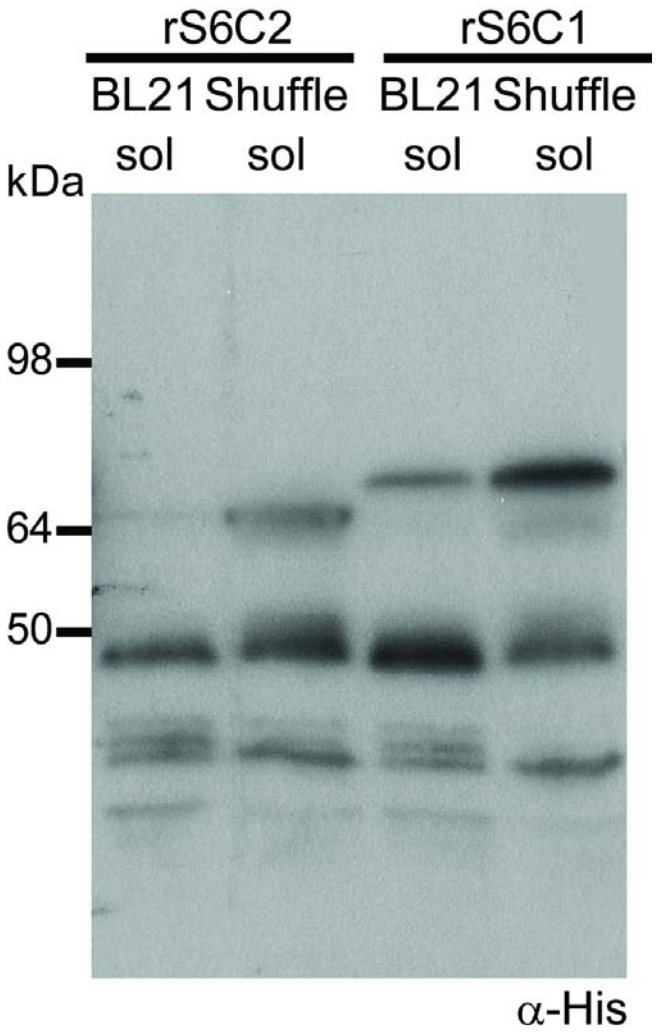
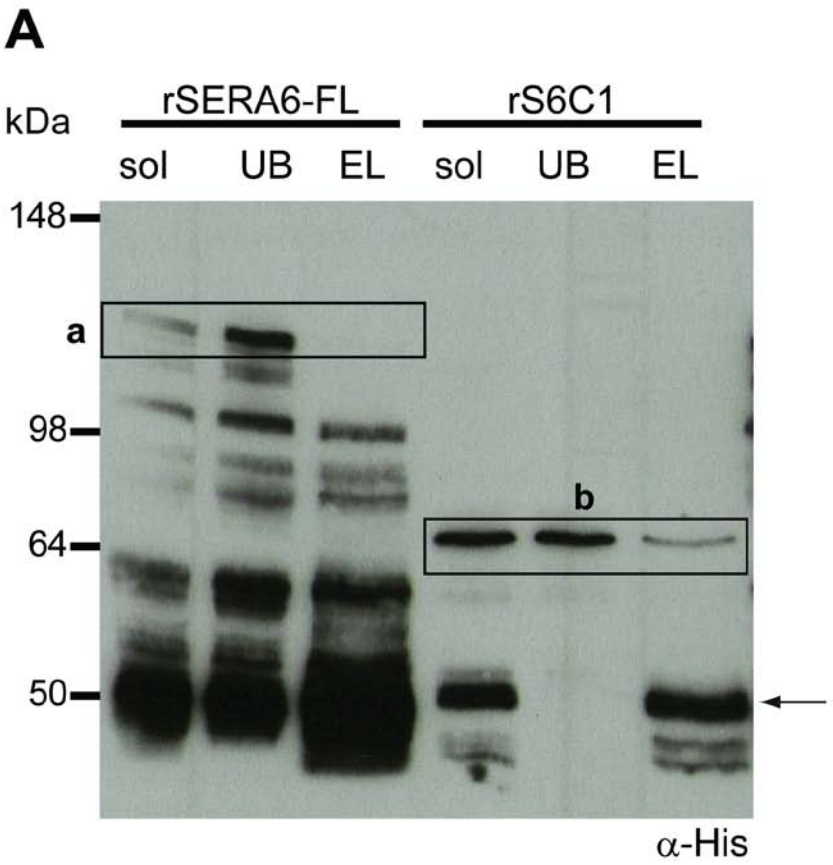

rS6C2
rS6C1
BL21
Shuffle
BL21
Shuffle
sol
sol
sol
sol
kDa
98
64
50
α-His

Figure 23: Purification of rS6C1 for antibody production

Recombinant S6-FL and rS6C1 were expressed in Shuffle T7 Express *lysY* cells overnight at 18°C. Cells were pelleted, solubilised in BugBuster Master Mix at RT and samples were subjected to SDS-PAGE. The majority of rS6-FL and rS6C1 appeared to not bind to the nickel resin and only a small amount was observed in the eluted fraction (A, a/b). The N-terminal tag containing breakdown products were copurified and constituted the largest protein quantity in the eluate (A, arrow). Despite numerous attempts, purification failed and focus was directed towards protein expressed in IB. All three SERA6 fragments were expressed in Shuffle T7 Express *lysY* cells at 18°C overnight and IB were washed several times followed by solubilisation in 8 M urea and fractions were separated by centrifugation. Very little recombinant protein was seen in the soluble fraction after BugBuster Master Mix lysis which was loaded as control (B, sol). The majority of rS6-FL and rS6C1 was detected in the urea fraction whereas only half of rS6C2 appeared to be solubilised (B, a-c). In 8 M urea solubilised rS6C1 was applied to a PD-10 desalting column and urea was exchanged for PBS and subjected to mice immunisation. sol: soluble fraction after BugBuster Master Mix lysis; UB: unbound after Nickel-Histidine chelation; EL: eluted fraction; urea: solubilised in urea; urea ins: insoluble after urea solubilisation


A
kDa
rSERA6-FL
rS6C1
sol
UB
EL
sol
UB
EL
148
98
64
50
a
b
α-His

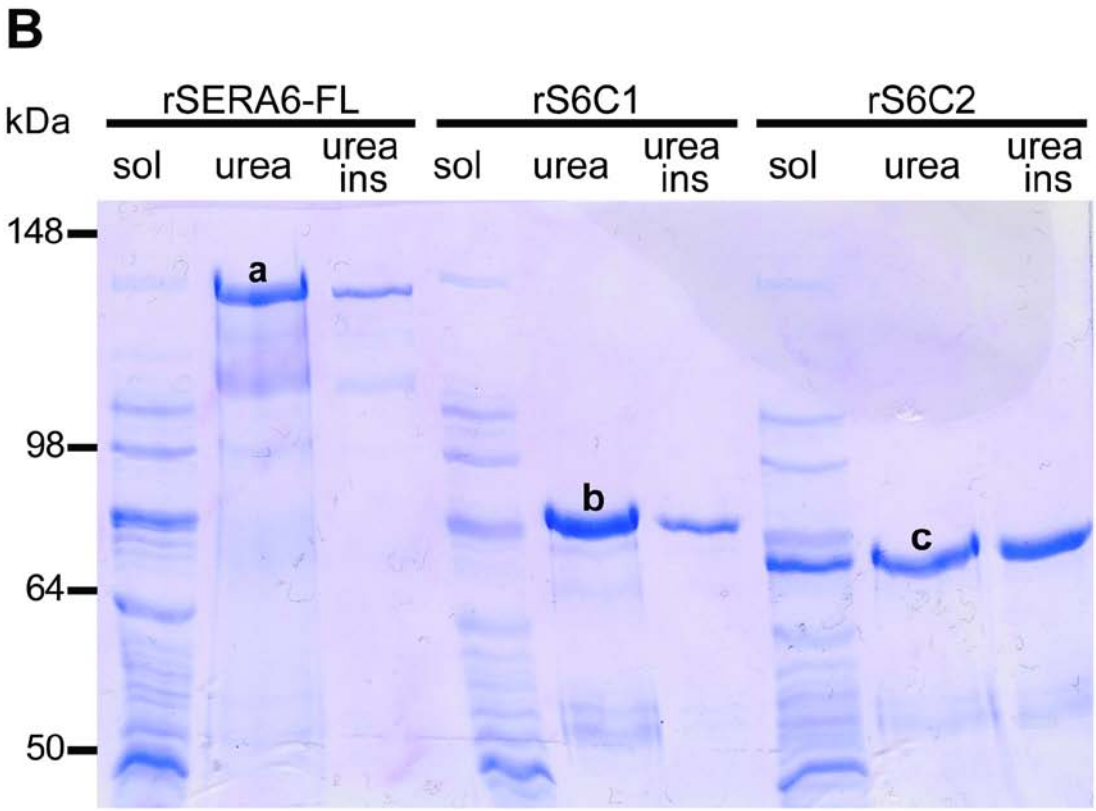

B
kDa
rSERA6-FL
rS6C1
rS6C2
sol
urea
urea ins
sol
urea
urea ins
sol
urea
urea ins
148
98
64
50
a
b
c

Figure 24: Recombinant expression of the SERA6 C-terminus leads to soluble protein and antibody production

A small C-terminal 6xHis- and S-tagged peptide was successfully expressed, purified and used for antibody production. The portion of SERA6 C-terminus used for recombinant expression was designed based on alignments with known PfSUB1 processing sites in SERA5. The soluble recombinant C-terminus (rS6-Cterm) was obtained by expression in Shuffle T7 Express *lysY* cells and purified using Nickel-Histidine chelation. All samples were separated via SDS-PAGE. The rS6-Cterm was successfully purified (A, asterisk) albeit not without the copurification of the previously described breakdown products (B, asterisk). Despite several attempts the separation of the rS6-Cterm and the breakdown products was unsuccessful and the eluted fraction was used in mice immunisations in order to raise antibodies. sol: soluble fraction after BugBuster Master Mix lysis; UB: unbound after Nickel-Histidine chelation; EL: eluted fraction; Western blot probed with α-His antibodies

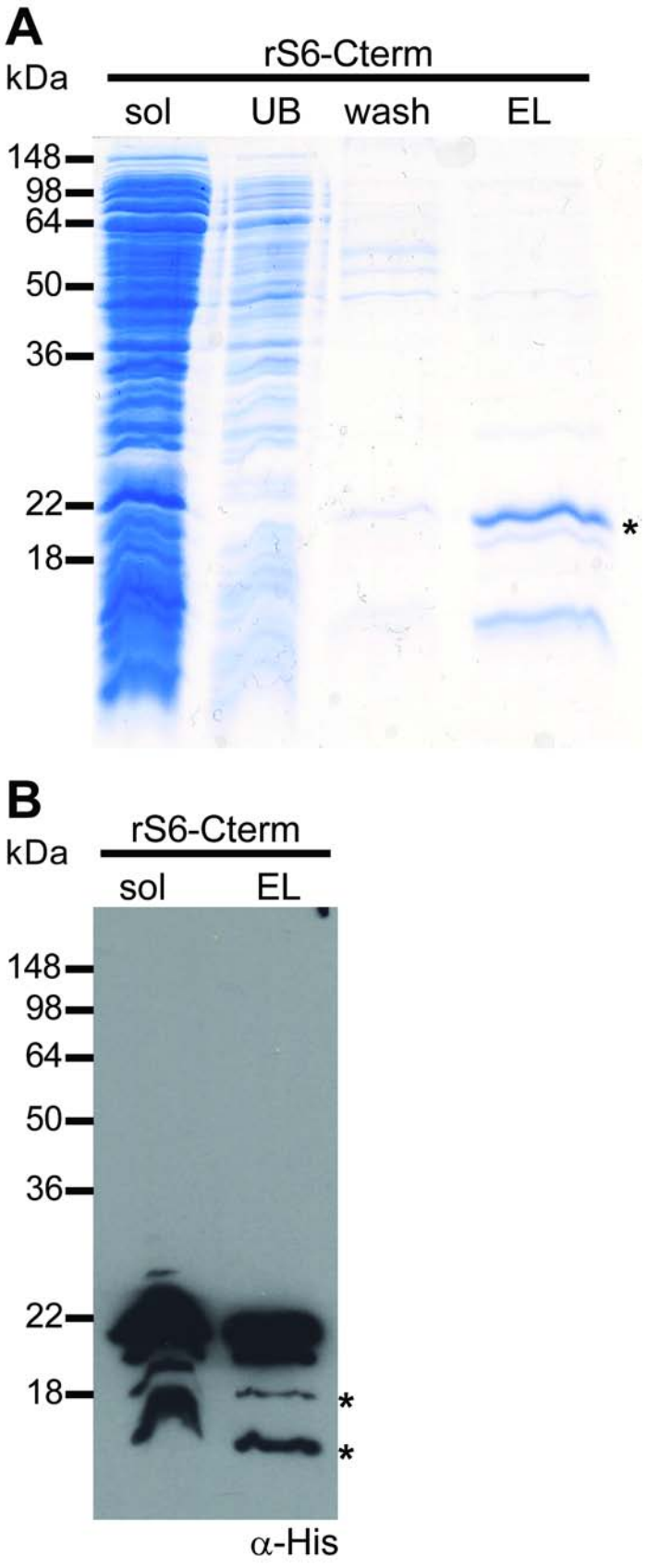

A
kDa
rS6-Cterm
sol
UB
wash
EL
148
98
64
50
36
22
18
*
B
kDa
rS6-Cterm
sol
EL
148
98
64
50
36
22
18
*
*
α-His

Figure 25: SERA6 is specifically recognised by antibodies raised against three different regions of the protein

Schematic of SERA6 protein, showing the putative catalytic residues (C644, H810, N835) in the shaded papain-like domain and the predicted SERA6st1 (VKAQ↓DDFN) and SERA6st2 (VHGQ↓SNES). The signal peptide is indicated by a black box. The three different antibodies were raised against the N-terminal region (α-S6-Nterm), the central region (α-S6C1) and the C-terminal region (α-S6-Cterm). The bars show the regions of SERA6 used for immunisation (A).

To test the specificity of the antibodies schizont lysates were probed by Western blot using monoclonal antibodies raised against rSERA5 (α-SERA5, mAb 24C6.1F1), rSERA6 central region (α-S6C1), rSERA6 C-terminus (α-S6-Cterm) or rSERA6 N-terminus (α-S6-Nterm, mAb 1C9). A single major band is observed at approximately 135 kDa for SERA6 and 120 kDa for SERA5. Membranes were subsequently stripped and reprobed with α-SERA5 or α-SERA6 to confirm that the antibodies did not cross-react. Faint bands at lower molecular weights may indicate early processing products of SERA5 (asterisks). downward-pointing arrow indicates site of cleavage

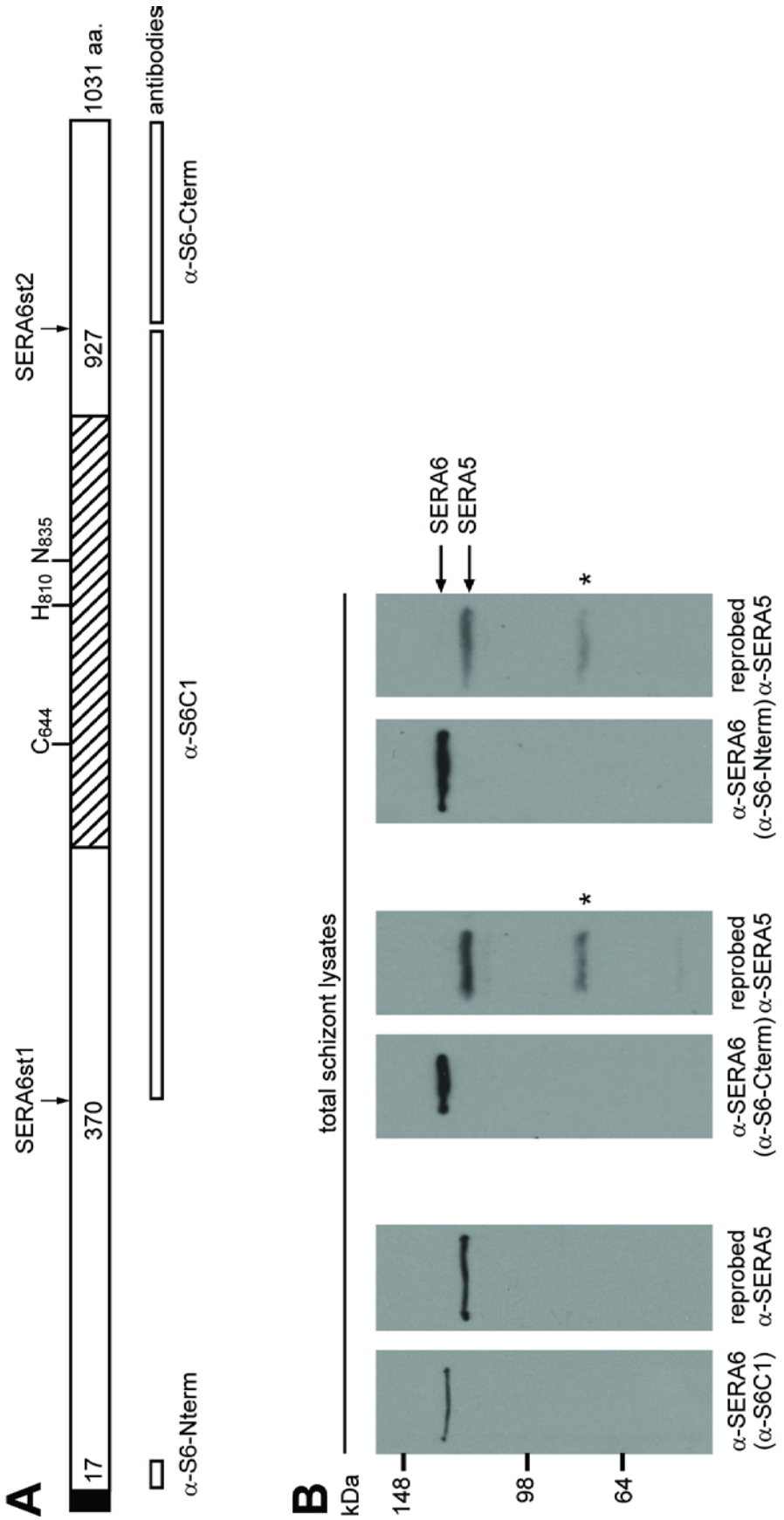

A
SERA6st1
SERA6st2
1031 aa.
17
370
927
C644
H810
N835
α-S6-Nterm
α-S6C1
α-S6-Cterm
antibodies
B
kDa
148
98
64
total schizont lysates
SERA6
SERA5
*
*
α-SERA6 (α-S6C1)
reprobed α-SERA5
α-SERA6 (α-S6-Cterm)
reprobed α-SERA5
α-SERA6 (α-S6-Nterm)
reprobed α-SERA5

Figure 26: Antibodies against SERA6 show low affinity in immunoprecipitation

Frozen schizont samples were hypotonically lysed and incubated with the monoclonal α-S6-Nterm antibody followed by incubation with Protein G beads. The unbound fraction was collected after centrifugation and the beads were several times washed in PBS followed by an elution step with 2x SDS sample buffer and Western blot analysis. SERA6 was detected in the elution fraction (A, probed with polyclonal rabbit α-S6C1). The nitrocellulose membrane was then stripped and reprobed with monoclonal α-SERA5 antibodies (mAb 24C6.1F1) to verify that SERA6 and not SERA5 was immunoprecipitated (B). The appearance of a band in the wash fractions and second band below SERA6 confirmed these findings.

In a next step it was attempted to establish appropriate conditions to separate the antibody and SERA6 in different elution fractions. Samples were prepared as mentioned above. The beads were pelleted, unbound material was collected (UB) the resin was washed using 1. PBS (wash1) followed by 50 mM Tris-HCl, 5 mM EDTA, 0.5% (v/v) Tx-100 pH 8.2 containing either no added salt (wash2), 500 mM NaCl (wash3) or 1 M NaCl (wash4). After the final wash beads were directly diluted in 2x SDS sample buffer. SERA6 can be detected in the 500 mM NaCl fraction without the presence of the antibody (C, probed with rabbit α-S6C1 antibody) as antibody can be found in the final fraction (D, EL, probed with mouse α-S6C1). Half of the SERA6 can be detected in the unbound fraction (C) which might me due to the lack of a sufficient amount of Protein G or insufficient binding between the antibody and the antigen. total: total schizont lysate; ST: start; UB: unbound; EL: elution in 2xSDS sample buffer

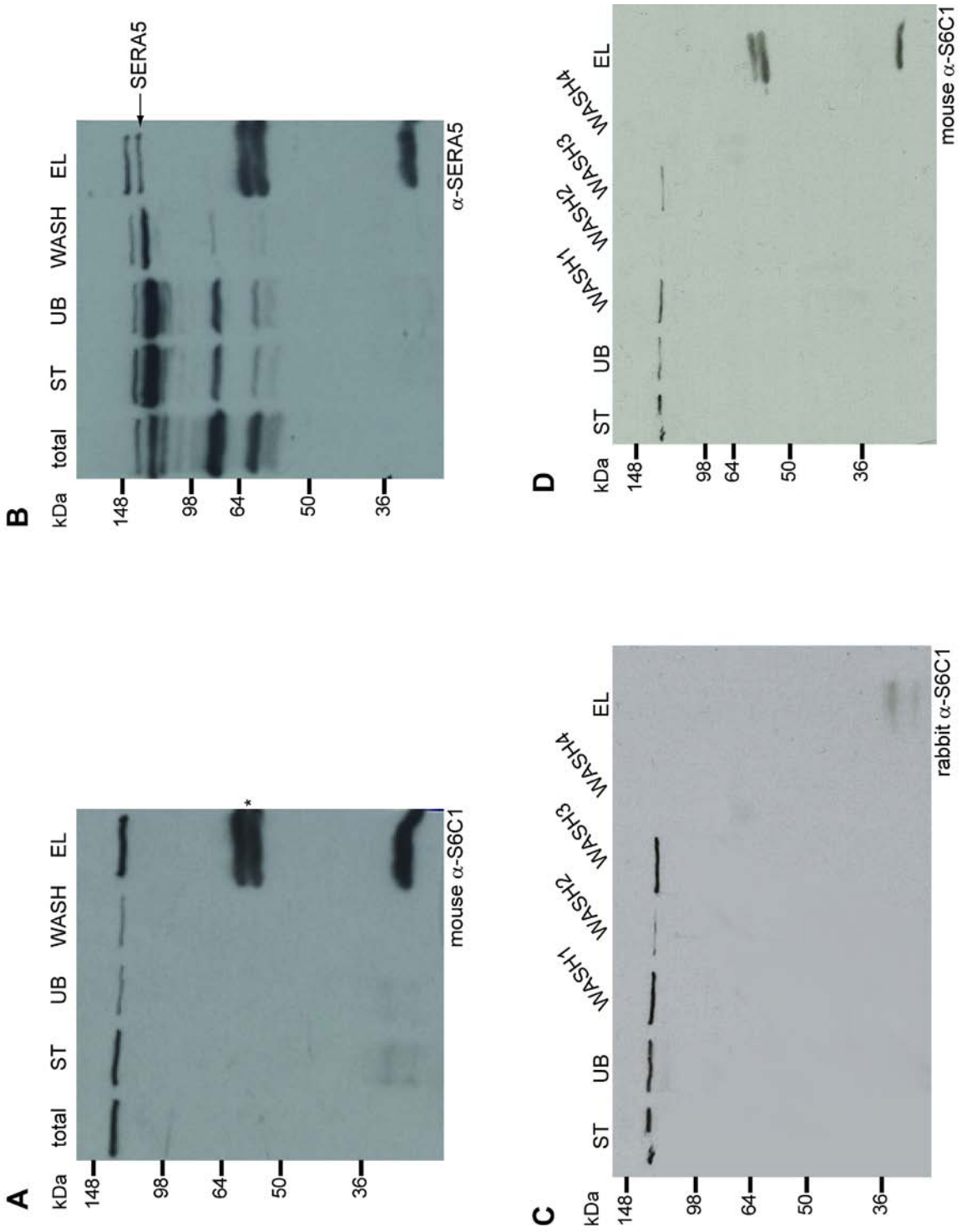

A
kDa
total
ST
UB
WASH
EL
148
98
64
50
36
*
mouse α-S6C1
B
kDa
total
ST
UB
WASH
EL
SERA5
148
98
64
50
36
α-SERA5
C
kDa
ST
UB
WASH1
WASH2
WASH3
WASH4
EL
148
98
64
50
36
rabbit α-S6C1
D
kDa
ST
UB
WASH1
WASH2
WASH3
WASH4
EL
148
98
64
50
36
mouse α-S6C1

Figure 27: Immunofluorescence microscopy suggests that SERA6 localises to the PV of schizonts

Fixed schizont smears were probed using polyclonal rabbit α-SERA6 (α-S6C1) antibodies (green), as well as monoclonal α-SERA5 (mAb 24C6.1F1, PV marker) or monoclonal α-MSP1 (mAb X509, labels the merozoite surface) antibodies (red). Nuclei were stained with DAPI. SERA6 partially showed a similar distribution with the PV marker SERA5 but it appeared as if it displayed a stronger staining at the periphery of the iRBC (A, late schizont and B, early schizont). The SERA6 signal appeared to surround the merozoites (α-MSP1, C-F). In some late schizonts, SERA6 staining was absent (D). This may be the result of diffusion of SERA6 out of the PV after permeabilisation or breakdown of the PVM during late schizont stages but with a still intact EPM. The observed staining patterns were not due to antibody effects as of two juxtaposed schizonts one was labelled and the other schizont was not (E). No signal against the central domain of SERA6 (α-S6C1) was observed on the surface of merozoites (F).

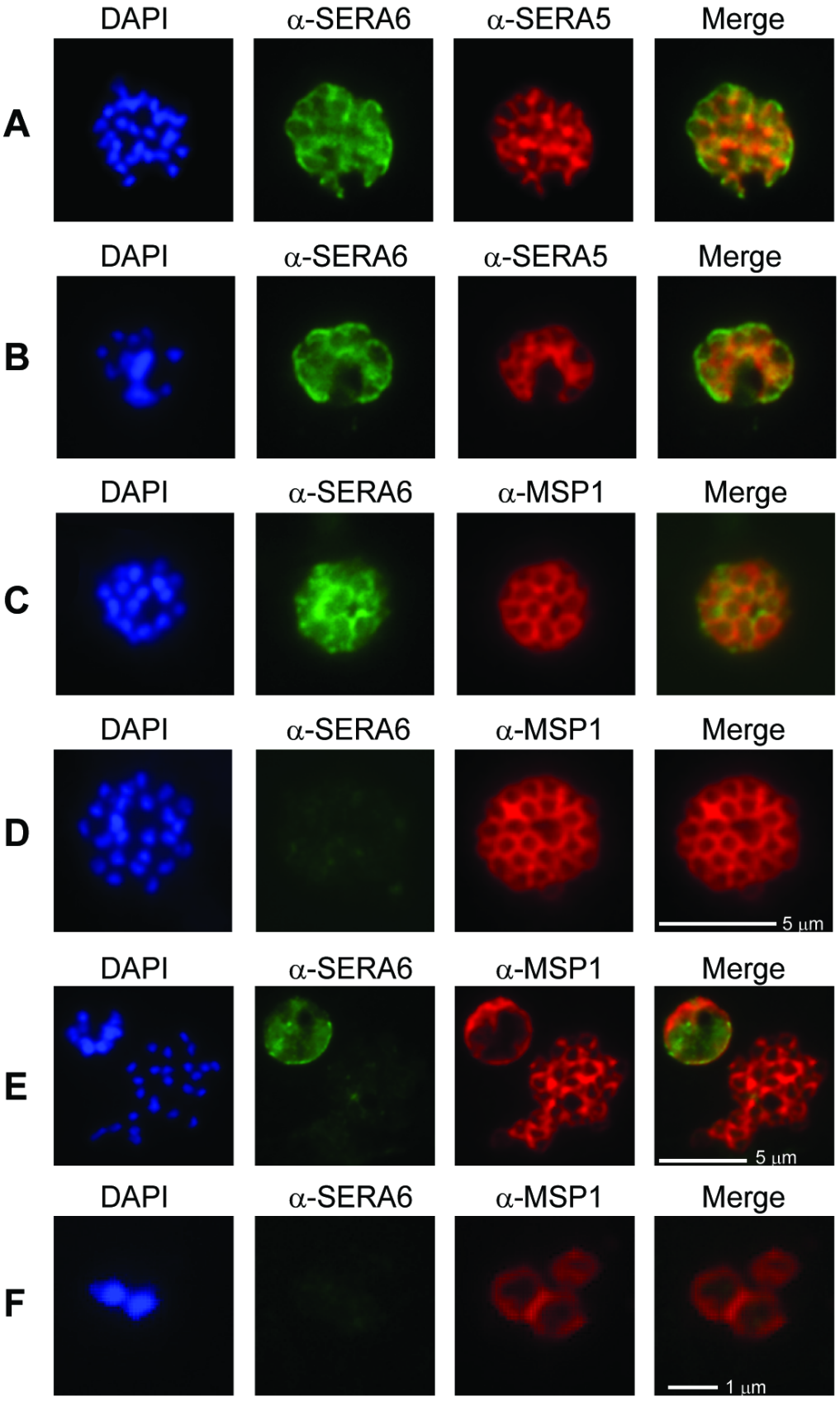

DAPI
α-SERA6
α-SERA5
Merge
A
DAPI
α-SERA6
α-SERA5
Merge
B
DAPI
α-SERA6
α-MSP1
Merge
C
DAPI
α-SERA6
α-MSP1
Merge
D
5 µm
DAPI
α-SERA6
α-MSP1
Merge
E
5 µm
DAPI
α-SERA6
α-MSP1
Merge
F
1 µm

Figure 28: The SERA6 *P. knowlesi* orthologue displays similar subcellular localisation

P. knowlesi parasites were grown in human blood and fixed schizont smears were probed using polyclonal rabbit α-S6C1 antibodies (green) as well as antibodies against *P. knowlesi* MSP19 which labelled the surface of merozoites (mouse polyclonal, red). Treatment with the SERA6 central antibody gave a similar signal distribution in *P. knowlesi* and the above described pattern of *P. falciparum* late stage parasites, with an accumulation towards the parasite periphery. The cross-reactivity of the antibody with *P. knowlesi* parasites was due to the fact that it was raised against the SERA6 central domain, which is highly conserved between both SERAs. The signal for the *P. knowlesi* SERA6 orthologue appeared to surround the merozoites with the same pattern as SERA6 but without display of overlapping signal distribution with PkMSP19.

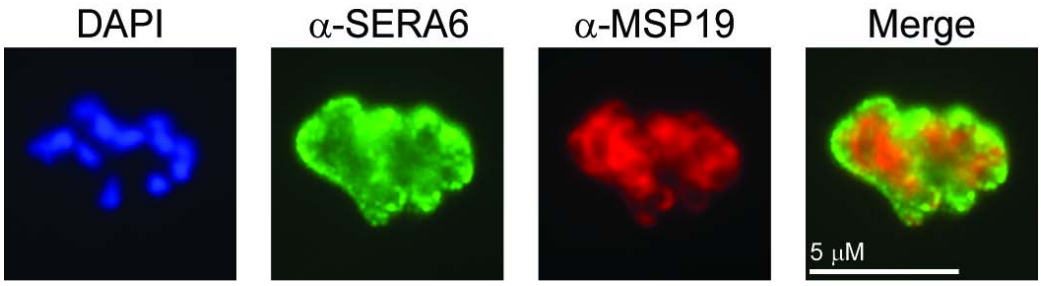
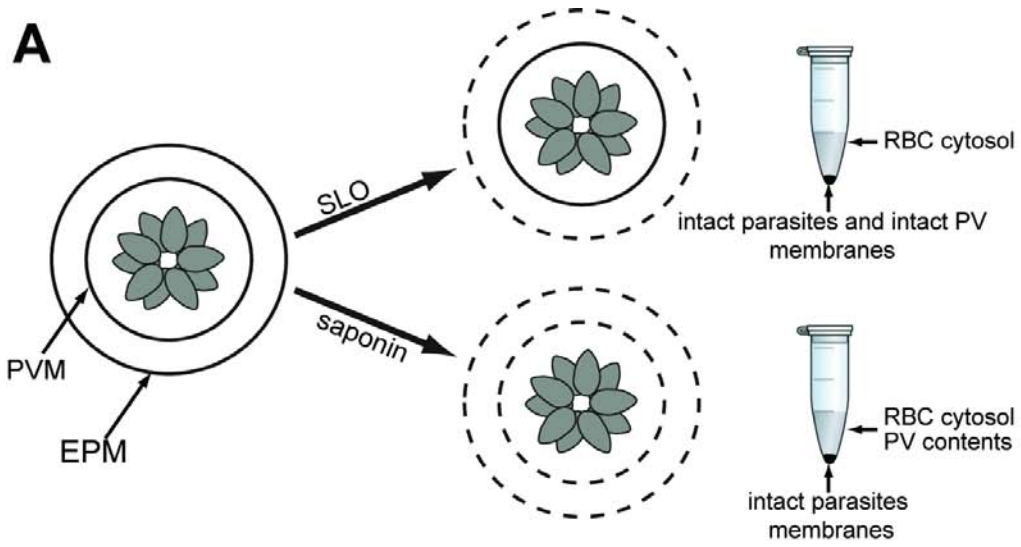
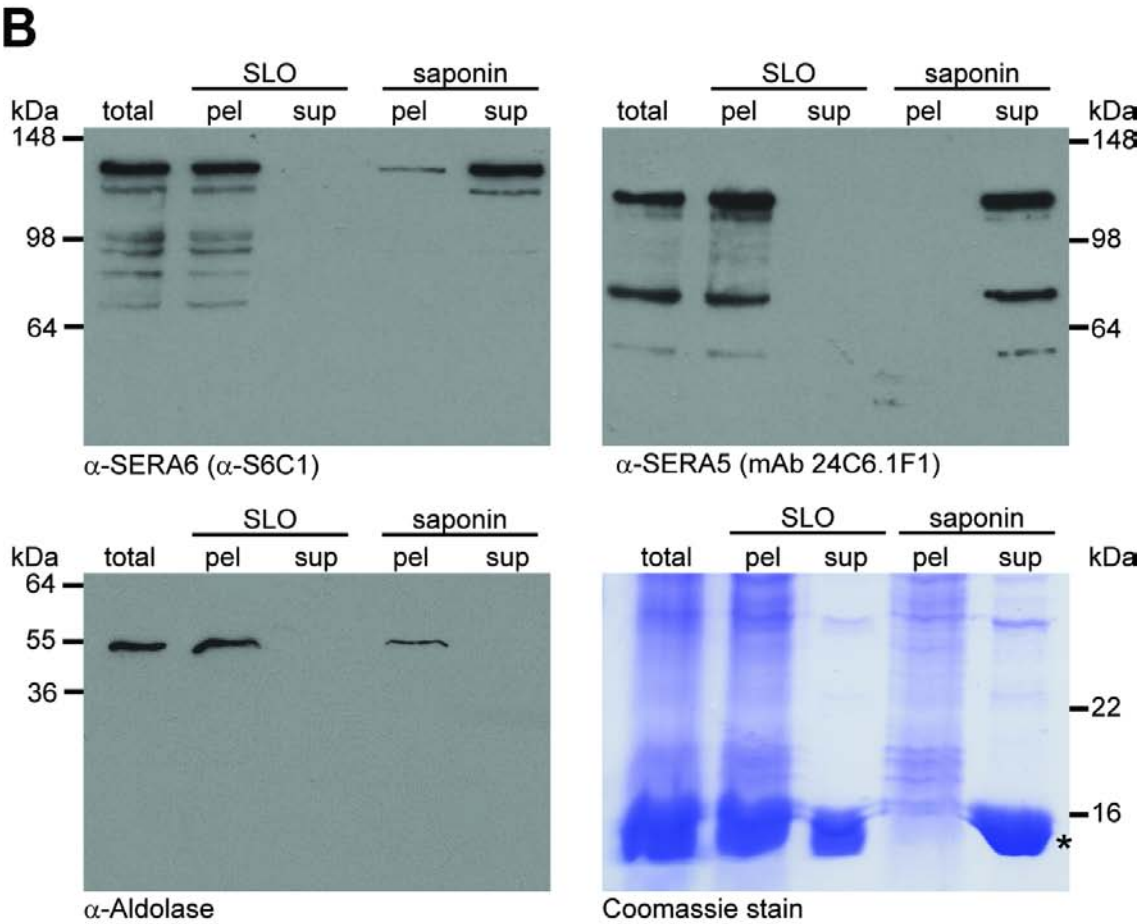

DAPI
α-SERA6
α-MSP19
Merge
5 μM

Figure 29: Biochemical fractionation of iRBCs suggests that SERA6 is a soluble protein localised to the PV

Overview of the methods used for biochemical fractionation of iRBCs (A). Upper half of schematic: SLO permeabilises the EPM leading to the release of soluble cytosolic iRBC material. Bottom half of schematic: saponin permeabilises both the EPM and PVM, leading to the release of soluble RBC and PV material.

Schizonts were fractionated with SLO or saponin. Pellet and supernatants were separated by SDS-PAGE and probed by Western blot using rabbit α-S6C1 (B, top left panel), α-SERA5 (B, top right panel) or α-Aldolase (B, bottom left panel). The samples were also stained with Coomassie Blue to reveal the major protein species at 16 kDa, haemoglobin (B, bottom right panel). SERA6 was present in the pellet fraction after SLO treatment (pel), but in the supernatant after saponin treatment (sup), suggesting that SERA6 is a soluble PV protein. A small amount of SERA6 remained in the saponin pellet fraction. SERA5, a known PV protein, served as a positive control. Aldolase, a cytosolic parasite protein, served as control for parasite lysis. Haemoglobin, a soluble RBC protein, served as a control for SLO and saponin permeabilisation of the RBC. EPM: erythrocyte plasma membrane; PVM: parasitophorous vacuole membrane; total: total schizont lysate; pel: pellet; sup: supernatant


A
SLO
saponin
PVM
EPM
RBC cytosol
intact parasites and intact PV membranes
RBC cytosol
PV contents
intact parasites membranes


B
SLO
saponin
kDa
total
pel
sup
pel
sup
148
98
64
α-SERA6 (α-S6C1)
α-SERA5 (mAb 24C6.1F1)
55
36
α-Aldolase
22
16
*
Coomassie stain

4 SERA6 processing by PfSUB1 is indispensable in the parasite

4.1 Introduction

The aims of the work described in this chapter were to examine the processing of SERA6 by PfSUB1 *in vitro* and in the parasite and to determine the physiological relevance of this processing. The experimental approaches were as follows: *in vitro* processing of SERA6 was first analysed by processing both rSERA6 derived from COS-7 cells and parasite-derived SERA6 with rPfSUB1. These results were then used to investigate SERA6 processing in the parasite. Next, identification of the PfSUB1 processing sites was attempted by N-terminal sequencing of processed rSERA6 (Edman, 1949). This was followed by addressing the question of whether SERA6 processing by PfSUB1 is important for parasite survival by attempting to modify processing sites in the genomic *sera6* gene. In a final step it was attempted to verify the conservation of PfSUB1 processing within SERA6 and its orthologue PbSERA3 to establish PbSERA3 as a model for SERA6.

4.2 Results

4.2.1 Processing of recombinant SERA6 by recombinant PfSUB1 generates fragments of predicted molecular masses

As recombinant expression of soluble and correctly folded SERA6 in bacteria proved challenging (see 3.2), extensive efforts were made to explore a variety of alternative expression strategies. COS-7 cells are derived from green monkey kidney cells and express the large T antigen of the simian vacuolating virus 40 (SV40) which increases replication and transcription efficiency of vectors carrying the SV40 promotor and origin of replication (Gluzman, 1981).

To attempt expression of SERA6 using the COS-7 cell system, the SERA6 WT coding sequence and rSERA6 C644A were cloned into the pSecTag_2A vector, in both cases fused to the plasmid-derived murine N-terminal Ig κ leader sequence to ensure secretion of the recombinant protein (Coloma *et al.*, 1992). The constructs were transfected into COS-7 cells and a variety of small scale expression conditions examined to identify the conditions that led to the highest amount of rSERA6 secreted into the culture supernatant. As it is unknown whether any *Plasmodium* proteins undergo N- or O- linked glycosylation in the parasite the transfected COS-7 cells were cultures in the presence of tunicamycin to prevent N-glycosylation during the expression of rSERA6 in COS-7 cells (Ulhoa *et al.*, 2001). Recombinant expression samples, including supernatants and cells, were collected at 2, 3 or 4 days after transfection and analysed by Western blot using rabbit α-S6C1. The same Western blot also included a schizont extract control, which was obtained by lysing late stage schizonts directly with SDS sample buffer. As shown in Figure 30, highest quantities of

soluble recombinant protein were obtained after transient expression for 4 days. However, Figure 30 also shows that the majority of rSERA6 accumulated intracellularly. Only a small proportion of the total rSERA6 was secreted into the culture supernatant. This was detected by Western blot examination of 4-day culture supernatants as a closely-spaced doublet. The lower band of this doublet comigrated with the full-length SERA6 signal detected in schizont samples (see Figure 30A, asterisk). The top band of the doublet was predicted to be glycosylated form of rSERA6, perhaps resulting from inefficient blockade of glycosylation by tunicamycin. To attempt to increase the efficiency of the tunicamycin treatment, it was therefore decided to examine whether an increase in the concentration of tunicamycin could completely prevent production of the higher molecular mass component of the doublet. As shown in Figure 30B, it was indeed found that increasing tunicamycin levels in the expression experiments resulted in secretion of only a single form of rSERA6 corresponding to the faster-migrating component of the doublet, consistent with it now efficiently blocking N-glycosylation of the secreted protein .

Having shown successful expression of soluble rSERA6, it was now possible to investigate its proteolytic processing by rPfSUB1. The rSERA6 containing COS-7 cell culture supernatants were concentrated, resuspended in PfSUB1 activity buffer and incubated in the presence of rPfSUB1. To ensure that any rSERA6 processing observed was solely due to rPfSUB1 and not contaminating proteases, samples of the rSERA6 were also incubated either alone or with both rPfSUB1 and rPD, a potent and selective inhibitor of PfSUB1. Mixtures were incubated at 37°C and samples taken at different time points were analysed by Western blot using the rabbit α-S6C1. As shown in Figure 31, these experiments showed that incubation of either rSERA6 WT or rSERA6 C644A with rPfSUB1 resulted in conversion of the presumed full-length 135 kDa rSERA6 to a ~75 kDa fragment (see Figure 31, c), via two intermediate products of ~125 and ~100 kDa (see Figure 31, a and b respectively). Minimal processing of SERA6 was observed in the additional presence of rPD, consistent with the observed processing being derived from rPfSUB1 activity, rather than from contaminating proteases; the very low level processing observed in the rPD sample is assumed to be due to an incorrect ratio of rPfSUB1 and the inhibitory rPD, under which conditions rPfSUB1 would be incompletely inhibited.

As described above, an *in silico* analysis of SERA6 cleavage sites predicted processed fragments of ~120 kDa (full-length), ~78 kDa and ~64 kDa. The ~78 kDa fragment was expected to be seen after processing at SERA6st1 and the ~64 kDa was predicted to appear after processing at SERA6st1 and SERAst2, both forms contain the central papain-like protease domain. As it is known that full-length SERA6 migrates slower than predicted on SDS-polyacrylamide gels under reducing conditions (Knapp

et al., 1991), the fragments observed at ~135 kDa, ~100 kDa and 75 kDa might therefore resemble the predicted SERA6 forms. In conclusion the results from this section demonstrated for the first time that rSERA6 can be a substrate for PfSUB1. The data also suggested that SERA6 might be processed at the predicted cleavage sites based on predicted molecular masses.

4.2.2 Parasite-derived SERA6 is processed by PfSUB1 *in vitro*

The results from the previous section showed that rSERA6 can be processed by rPfSUB1 *in vitro*, leading to three different fragments detectable by Western blot with antibodies specific for α-S6C1. The next set of experiments aimed to examine the important question of whether SERA6 processing by PfSUB1 also occurs in the parasite. This was approached in two ways - firstly, by analysing the processing of parasite-derived SERA6 *in vitro*; and secondly, by comparing these processed samples with mature schizont extracts on Western blots, harvested under conditions where PfSUB1 discharge into the PV would likely have taken place at least to some extent, potentially exposing the endogenous SERA6 to the protease.

Before processing of parasite-derived SERA6 could be analysed *in vitro* the protein had to be purified from parasite extracts. Synchronous late stage 3D7 parasites (harvested 42 hours post-invasion) were hypotonically lysed and the extracts submitted to size-exclusion chromatography followed by anion-exchange chromatography, as previously described for SERA5 (Yeoh *et al.*, 2007). As seen in Figure 32 SERA6 was successfully purified. The anion-exchange chromatography peak fractions (fractions 33 and 34) that contained parasite-derived SERA6 were then pooled, concentrated and resuspended in PfSUB1 activity buffer. The preparations are subsequently referred to as parasite-derived SERA6.

To address the question of whether parasite-derived SERA6 can be processed by rPfSUB1 *in vitro*, samples of parasite-derived SERA6 were then incubated with rPfSUB1. To ensure that any processing seen was solely due to rPfSUB1 and not contaminating proteases, possibly copurified with parasite-derived SERA6, samples of parasite-derived SERA6 were also incubated alone or with both rPfSUB1 and rPD. Samples of the digests were taken at different time points and analysed by Western blot using the rabbit α-S6C1. Figure 33A shows that processing by rPfSUB1 was detectable after 30 min, whereas the control preparations showed no or very little proteolysis. The SERA6 processing seen was not caused by contaminating proteases, as samples lacking rPfSUB1 were stable even after two hours at 37°C. The sample containing both rPfSUB1 and rPD showed slight processing, but this was probably the result of adding the enzyme moments before the rPD. Processing appeared not to go to completion in the presence of rPfSUB1 as even after two hours several fragments

could be observed. The results demonstrated that parasite-derived SERA6 was processed by rPfSUB1. The processing fragments observed at ~125, ~100 kDa and ~75 kDa appear identical to the processing fragments of COS-7 cell derived SERA6.

4.2.3 PfSUB1 is responsible for SERA6 processing in the parasite

Although the above experiments showed that parasite-derived SERA6 can be processed *in vitro* by rPfSUB1, the important question remained whether this also occurred under endogenous conditions in the parasite. It was therefore decided to assay whether the pattern of processing that appeared during SERA6 processing *in vitro* can also be observed in very late stage 3D7 parasites where it was anticipated that release of PfSUB1 into the PV would have commenced. If PfSUB1 was responsible for processing of endogenous SERA6 in the parasite, it would be predicted that SERA6 fragments from late stage schizonts and the *in vitro* processed parasite-derived SERA6 would comigrate on Western blot. Late stage schizonts were obtained by monitoring synchronised parasite cultures from 36 hours post invasion onwards assessing the maturity of the schizonts by examination of Giemsa stained smears every 60 min. At the first signs of schizont rupture, the parasite cultures were collected, pelleted and extracted in sample buffer to produce schizont extracts. These were then examined by Western blot in comparison with *in vitro*-processed parasite-derived SERA6. Figure 33B shows that the *in vitro* processing fragments and fragments from the processing in the parasite appeared to comigrate at least partially. This suggests that the processing of SERA6 observed *in vitro* mimics endogenous SERA6 processing in the parasite.

The results shown above were the first ones to imply that endogenous SERA6 is indeed processed by PfSUB1 in the parasite. To further explore whether *in vitro* processing of SERA6 is similar to the processing pattern of endogenous SERA6 in the parasite, further assays were done to examine whether products of COS-7 cell derived rSERA6 processed by rPfSUB1 comigrate with endogenous SERA6 fragments from schizont extracts. COS-7 cell derived rSERA6 was processed with rPfSUB1 and samples taken at four different time points were compared to schizont extracts by Western blot. As shown in Figure 33C all the bands corresponding to processed fragments of rSERA6 comigrate with bands visible in schizont extracts. This suggested that processing of COS-7 cell derived SERA6 *in vitro* was indeed similar to endogenous SERA6 processing in the parasite. In conclusion the assays performed in this section have shown that PfSUB1 appears to process SERA6 in the parasite.

4.2.4 Confirmation of previously predicted PfSUB1 cleavage sites and identification of an additional and previously unpredicted N-terminal processing site

Having confirmed that endogenous SERA6 is indeed cleaved by PfSUB1 *in vitro* and in the parasite it was important to establish the precise identity of the PfSUB1 cleavage sites in SERA6. It was also noted that the appearance of a SERA6 fragment at ~125 kDa (just below full-length SERA6) during all processing assays indicated that SERA6 may contain a third PfSUB1 processing site. N-terminal sequencing is the only direct way to precisely determine protease cleavage sites after proteolytic processing. Therefore it was attempted to N-terminally sequence the fragments of rSERA6 produced following digestion with rPfSUB1 to identify all PfSUB1 processing sites. In anticipation that cleavage occurred at the predicted sites, a number of rSERA6 mutants were designed containing mutations expected to block PfSUB1 processing. This had the advantage that the cleavage sites were not only confirmed by N-terminal processing but it could also be shown that the alteration of the predicted cleavage sites prevented PfSUB1 processing

PfSUB1 has previously been shown to be incapable of cleaving peptides with a Leu residue in the P2 or P1 position (Koussis *et al.*, 2009). Accordingly, the following mutations (see Figure 34A) were introduced into the predicted PfSUB1 cleavage sites in SERA6: SERA6st1 (VKAQ↓DDFN to VK**LL**DDFN) to generate rS6-FLmutS1; SERA6st2 (VHGQ↓SNES to VH**LL**SNES) to generate rS6-FLmutS2 or both (rS6-FLmutS1S2). As rS6-FLmutS1S2 had both PfSUB1 processing sites altered, this recombinant protein was not expected to undergo any processing at all by rPfSUB1. All three mutants and SERA6 WT were expressed as IB in *E. coli*, and the proteins purified by excision from an SDS-polyacrylamide gel. The purified proteins were then in-gel digested with rPfSUB1, the digestion products separated by SDS-PAGE, transferred onto a PVDF membrane (suitable for N-terminal processing) and stained with Coomassie Blue (see Figure 34B for an experimental overview of this protocol).

Analysis of the Coomassie Blue stained SDS-polyacrylamide gels showed that the SERA6 WT (~135 kDa) was digested to a terminal product of ~75 kDa (seeFigure 35A, a). Recombinant S6-FLmutS1 was cleaved to form a larger final product of 110 kDa via an intermediate of 125 kDa (see Figure 35A, b). Recombinant S6-FLmutS2 was cleaved to a final product of 100 kDa (see Figure 35A, c). Surprisingly rS6-FLmutS1S2 (which was predicted to not undergo any processing) was also processed to a ~125 kDa product, indicating the presence of a third unpredicted PfSUB1 processing site (see Figure 35A, d). A small 14 kDa processing fragment, possibly representing the C-terminus of rSERA6, was only present in digests of rS6-FL or rS6-FLmutS1, (see Figure 35A, bottom panel, e). The results suggested that alterations

made to the PfSUB1 cleavage sites in SERA6 indeed affected processing in those sites.

N-terminal sequencing was then performed on the various digestion products to determine the sequence of the PfSUB1 cleavage sites. The results of these are shown in Figure 35B. The two previously predicted sites were confirmed, and one extra cleavage site in the N-terminus (SERA6st3) identified as VVSS↓SESG (see Figure 35B). In conclusion results from this section have shown that SERA6 is processed by PfSUB1 in three, and not as predicted, two processing sites. These results were verified by showing that the alteration of these sites prevent PfSUB1 processing of SERA6 *in vitro*.

4.2.5 Processing of SERA6 by PfSUB1 at at least one site is indispensable for parasite survival

Having shown that SERA6 is processed by PfSUB1 in the parasite, it was decided to address the biological importance of this by attempting to selectively perturb processing of endogenous SERA6 in the parasite. The results from the previous section showed that the modification of the PfSUB1 processing sites in SERA6 blocked processing of the recombinant protein *in vitro*. Therefore to prevent PfSUB1 processing in the endogenous SERA6, the above identified SERA6 cleavage sites were partially replaced or deleted in the genomic *sera6* gene in the 3D7 *P. falciparum* clone using single homologous recombination. If the maturation of SERA6 was required for parasite survival the integration of modified processing sites would not be tolerated and no homologous recombination events would be detected in those parasite lines. If however the processing at one or several sites was dispensable the respective parasite lines should integrate the modified processing sites and homologous recombination events would be detected.

To perform the targeted recombination experiments, a transfection plasmid was designed based on previously used strategies (Yeoh *et al.*, 2007;Child *et al.*, 2010). This involves production of a chimeric gene consisting of an N-terminal genomic segment of the *sera6* gene fused in frame to the C-terminal domain of a recodonised *sera6* gene. Integration of this construct into the *sera6* locus would be expected to reconstitute the entire *sera6* coding sequence in the form of a chimeric gene. Given that the recodonised *sera6* gene shares only little nucleotide identity with the genomic *sera6*, no recombination should occur downstream of the fusion between the genomic and the recodonised *sera6*, enabling the incorporation of desired mutations into the recodonised sequence. This approach also conferred advantage for the subsequent analysis of integrated lines. Nine integration constructs were produced, each containing mutations designed to alter the amino acid sequence of the PfSUB1 processing sites,

either individually or in combination. One additional construct (SERA6chim) was designed to leave the processing sites unaltered, thereby serving as a control to show that the *sera6* locus is accessible to homologous recombination. See Figure 36A for an overview of the construct design, whilst Table 6 shows an overview of all transfection constructs.

Each plasmid was transfected into 3D7 *P. falciparum* parasites on at least 3 independent occasions and the parasite cultures subjected to drug cycling to select for parasites in which integration has taken place. After two to four drug cycles, genomic DNA from the parasite lines was examined using a diagnostic PCR to test whether transfected plasmids had integrated as expected into the targeted *sera6* locus. The primers used for the diagnostic PCR were designed to bind to recodonised *sera6* and were unlikely to anneal to genomic *sera6* making the identification of integrant lines more specific (see Figure 36A).

According to the results from the PCR (see Figure 36B), integration was detected for: the SERA6 WT (pHH-SERA6chim) as well as SERA6st2 deletion (pHH-SERA6chim∆S2, -SNES), SERA6st2 partial substitution (pHH-SERA6chim_mutS2, SERA6st2: VH**LL**SNES) and the combination of SERA6st2 partial replacement and SERA6stX (pHH-SERA6chim_mutS2Sx, SERA6st2: VH**LL**SNES, SERA6stX: V**AAA**RHNP). Southern blot analysis was then performed in an attempt to confirm the integration of the PCR positive parasite lines. As shown in Figure 36C one band at ~3.3 kbp which is specific for a successfully integrated *sera6* locus was detected by Southern blot analysis for: pHH-SERA6chim, pHH-SERA6chim∆S2, pHH-SERA6chim_mutS2 and pHH-SERA6chim_mutS2Sx. The same Figure also showed a band which was indicative for the presence of the transfection plasmid or the genomic and unmodified *sera6* gene as both would be detected at ~9.5 kbp (see Figure 36C, asterisk).

Collectively, the results of this analysis showed that SERA6st1 appeared to be completely refractory to genetic disruption whereas SERA6st2 deletion and partial substitution were tolerated by the parasite. SERA6stX was refractory to genetic manipulation except in combination with the SERA6st2 substitution, which is surprising as SERA6stX was not integrated individually. Therefore some but not all PfSUB1 processing sites appeared to be required for parasite survival.

4.2.6 Recombinant PbSERA3 serves as a substrate for PfSUB1

4.2.6.1 Recombinant PbSERA3 undergoes processing by PfSUB1

Expression of *P. falciparum* SERA6 has proven to be difficult. In order to further explore the function of this protein it was decided to attempt expression of the *P. berghei* orthologue, PbSERA3. Recombinant PbSERA3 could not only give further

insights into the proteolytic maturation pattern of the “Cys-type” SERA family members, it might also be used to screen for proteolytic activity. Identifying the proteolytic activity of rPbSERA3 will beneficial for identifying a variety of substrates. These substrates could then be used to screen for the proteolytic activity of SERA6.

Alignments between PbSERA3 and SERA6 revealed that both carry PfSUB1 processing sites in similar locations (see Figure 15). This suggested that PbSERA3 was likely to undergo processing by PfSUB1 at two sites similar to SERA6 (SERA6st1 and SERA6st2) which will be referred to as PbSERA3st1 and PbSERA3st2. Two different PbSERA3 constructs for recombinant expression in insect cells were designed: One construct reconstituted the full-length PbSERA3 wild-type (called rPbSERA3 WT); the second construct had the putatively catalytic Cys residue substituted with an Ala residue creating a catalytic site mutant (called rPbSERA3 C639A). The rPbSERA3 C639A mutant was aimed to be used in activity studies of rPbSERA3 which will be described and discussed in chapter 5. Recombinant PbSERA3 WT and rPbSERA3 C639A were transiently expressed in insect cells (Tn5 cells) in the presence of tunicamycin and supernatants containing the recombinant protein were collected after three days. The recombinant protein was purified using anion-exchange chromatography followed by size-exclusion chromatography and eluted in PfSUB1 activity buffer. Fractions were analysed by Western Blot with α-PbS3C1 antibodies (against the central PbSERA3 domain). Figure 37A showed that both recombinant proteins were eluted between fraction 31 and 33 on the final chromatography step (see Figure 37A).

To next examine processing of PbSERA3 by PfSUB1, the purified proteins were subjected to processing assays with rPfSUB1, following a procedure similar to that used to examine recombinant and parasite-derived SERA6 (see 4.2.1). To ensure that any processing observed was due to rPfSUB1 and not contaminating proteases, samples of the rPbSERA3 were also incubated alone or in the presence of both rPfSUB1 and rPD. Figure 37B shows that both rPbSERA3 proteins were processed by rPfSUB1, with the full-length rPbSERA3 being converted efficiently to a fragment at ~100 kDa followed by a further conversion to fragments of ~75 kDa and ~48 kDa (see Figure 37B, arrows a-c). Interestingly, rPbSERA3 WT but not rPbSERA3 C639A were further converted to a fragment of ~42 kDa (see Figure 37B, arrows d-e). It was suggested that the 75 kDa fragment might correspond to the same fragment in SERA6, S6C1, and will now be referred to as PbS3C1. The results showed that both proteins are substrates for PfSUB1, although with a processing pattern that was not consistent with the predicted pattern.

4.2.6.2 Recombinant PbSERA3 undergoes processing by PfSUB1 in an additional and unpredicted internal processing site

The results of the previous section implied that rPbSERA3 WT and rPbSERA3 C639A are both processed by rPfSUB1. Only two PfSUB1 cleavage sites were originally predicted within PbSERA3, but the results showed that both rPbSERA3 proteins were processed at more than two sites as several fragments appeared during the rPfSUB1 processing assay. As rPbSERA3 should serve as a model for SERA6 it was important to establish a clear understanding of the processing of rPbSERA3. Therefore the next set of experiments attempted to identify the protease responsible for the processing step leading to the 48 kDa fragment. The appearance of a 42 kDa rPbSERA3 WT form will be discussed in chapter 5. The 48 kDa fragment was found to contain the central-papain like domain as it was repeatedly detected with antibodies which were raised against this central PbSERA3 domain. It was therefore likely that the 48 kDa fragment was the product of an additional processing of PbS3C1 (PbSERA3 between both predicted PfSUB1 cleavage sites). To test whether rPfSUB1 was responsible for this, a construct for the recombinant expression of PbS3C1 was designed. The resulting protein would precisely span the PbSERA3 sequence between the predicted PbSERA3st1 and PbSERA3st2 and it would also carry an N-terminal 6xHis-tag followed by an S-tag (see Figure 38A). It was expected that if the 48 kDa fragment was the product after processing rPbSERA3 WT between PbSERAst1 and PbSERA3st2, rPbS3C1 should be processed by PfSUB1 to a product of the same molecular mass as the 48 kDa fragment. The recombinant rPbS3C1 was expressed in *E. coli* where it accumulated in insoluble IB, the IB were solubilised and rPbS3C1 purified by Nickel-Histidine chelation. The rPbS3C1 was then incubated with rPfSUB1 at 37°C, samples were taken at two time points and analysed by Western blot with α-PbS3C1 antibodies. As seen in Figure 38B the unprocessed rPbS3C1 was detected at ~66 kDa. Samples containing rPbS3C1 and rPfSUB1 showed a major fragment appearing at ~48 kDa and the 66 kDa band was detected with only a very faint signal (see Figure 38B, left hand panel). Further examination of the same samples showed that the processed rPbS3C1 no longer reacted with the 6xHis-specific monoclonal antibody. This was due to the loss of the N-terminal 6xHis-tag suggesting that the processing took place at the N-terminal end of the rPbS3C1 (see Figure 38B, right hand panel). The results demonstrated that rPbS3C1 is cleaved by rPfSUB1 causing the appearance of a 48 kDa fragment.

To attempt to identify this third PfSUB1 processing site, samples of rPbS3C1 digested with rPfSUB1 were examined by N-terminal sequence analysis. The N-terminal sequence of the ~48 kDa digestion product was identified (see Figure 38C) consistent with cleavage taking place at the motif AIGQ↓NEEP (the downward-pointing

arrow indicates the site of cleavage). In summary, this analysis established that PbSERA3 undergoes proteolytic processing by PfSUB1 at at least three sites.

4.3 Discussion

4.3.1 Challenges of recombinant SERA6 expression

Recombinant protein expression is often accompanied by multiple challenges especially in proteins as cysteine rich as SERA6. Throughout this study six heterologous expression systems were tested to obtain soluble and correctly refolded SERA6. Most approaches led to insoluble protein or no protein expression at all. The most promising strategies turned out to be expression in bacterial cells and COS-7 cells. Bacterial cells produced very high protein quantities which were however insoluble. Extensive investigation into improvement of the bacterial expression conditions did not lead to soluble SERA6 and most purification attempts were unsuccessful. COS-7 cells secreted soluble full-length SERA6 into the culture supernatant but in very low quantities. Considerable efforts were made towards the improvement of expression conditions in COS-7 cells in order to increase soluble protein amounts. As the transfection efficiency was rather low (5%, the cells were transfected at three independent occasions) a variety of transfection methods were applied, e.g. electroporation or different concentrations of the transfection reagent. It is known that alteration of these conditions can have a huge impact on protein solubility and protein quantities. However, the increase to a transfection efficiency of 50% (the cells were transfected at three independent occasions) did not lead to an increase in recombinant protein amounts or elevated secretion levels. Despite these attempts to improve secretion levels the majority of the recombinant full-length SERA6 protein remained intracellular. As the COS-7 cells were grown attached to sterile culture flasks the transfection efficiency in future experiments may be increased by using cells cultured in suspension. This would increase the access to the cell surface area to internalise the plasmid vector.

4.3.2 SERA6 is processed by PfSUB1 *in vitro* and the parasite

The results from this chapter suggested that recombinant and parasite-derived SERA6 both undergo cleavage by rPfSUB1 *in vitro*. Further work showed that SERA6 processing patterns *in vitro* and in the parasite were mostly identical, strongly suggesting that PfSUB1 is responsible for cleaving SERA6 in the parasite. At the beginning of this work it was predicted that SERA6 undergoes processing by rPfSUB1 in two sites. Cleavage at these two steps would be expected to convert the full-length SERA6 (120 kDa) via an intermediate of ~90 or ~76 kDa (depending on the order of PfSUB1 cleavage) to a final fragment of ~64 kDa which contains the putative catalytic

domain. The experimental results from this chapter indicate that the three major fragments observed during PfSUB1 processing are ~125 kDa, ~100 kDa and ~75 kDa (based on migration on SDS-polyacrylamide gels). Full-length SERA6 migrated at ~135 kDa, larger than the predicted 120 kDa. The appearance of a band at ~125 kDa was not predicted as an event caused by PfSUB1. This suggested that SERA6 contains an additional PfSUB1 cleavage site on the N-terminus, similar to that observed for some allelic forms of SERA5 (Li *et al.*, 2002b). It can be excluded that this step might have been an *in vitro* processing artefact as bands of the same molecular weight appeared during processing in the parasite. This event is likely to be the first SERA6 cleavage step as the 125 kDa fragment disappeared during processing assays and smaller products accumulated. Results from N-terminal sequencing revealed that the putative catalytic fragment of 64 kDa may correspond to the observed 75 kDa band. Therefore it can be proposed that full-length SERA6 (135 kDa) is processed by PfSUB1 at SERA6st3 to a fragment of 125 kDa. This is then followed by a second processing step in SERA6st1 to a fragment at 100 kDa. The PfSUB1 processing is then finalised after the cleavage at SERA6st2 to release a 75 kDa fragment (see Figure 39).

In the parasite, it is likely that SERA6 could also undergo further processing by different proteases or through autocatalytic processing. This is supported by findings that in some processing assays endogenous SERA6 was additionally processed to several smaller fragments. SERA5 also undergoes further processing from P56 to P50 which is caused by a leupeptin sensitive cysteine protease called Protease X (Sharon Yeoh and Michael Blackman, unpublished). It may be that SERA6 also undergoes processing by Protease X.

The data from this chapter suggests that SERA6 processing occurs very late in the parasite life-cycle, only moments before egress. This is supported by the findings that continuous monitoring of late stage parasite cultures by Giemsa staining and Western blot analysis only showed first signs of SERA6 processing once the first merozoites started to escape.

4.3.3 Partial SERA6 processing by PfSUB1 is crucial for parasite survival

Work in this chapter demonstrated that SERA6 is processed by PfSUB1 at three sites. As SERA6 was previously shown to be indispensable for the parasite (Miller *et al.*, 2002b) the question was asked whether the proteolytic maturation of SERA6 was also required for parasite survival *in vitro*. Therefore it was aimed to modify the endogenous SERA6st1, SERA6st2 and SERA6stX individually or mixed in all possible combinations. The results suggested that the processing of SERA6st1, which is located upstream of the central papain-like protease domain, appeared to be important independently whether transfected individually or in combination. The modification of

SERA6st2 appeared to be tolerated by the transfected parasite lines. Whether modifications of SERA6stX are detrimental to the parasite is unknown as the alteration of this site appeared to be tolerated in combination with modifications to SERA6st2 but not if SERA6stX was altered individually. The parasite lines which appeared to tolerate blocking of PfSUB1 processing have yet to be cloned. This would then allow for the analysis of the effect of the processing site alterations; particularly to answer the question of whether processing by PfSUB1 is in fact inhibited. This is important as similar transfection approaches in SERA5 showed that SERA5 still underwent processing in cloned SERA5st2 mutant parasite lines (Christine Collins, unpublished). It therefore may be that SERA6st2 mutant lines still undergo processing by PfSUB1. If, however, SERA6st2 is not processed but parasites survive, this would indicate that processing at SERA6st1 alone is sufficient for SERA6 to function normally. Interestingly it was also shown that the alteration of the processing sites *in vitro* completely blocked the cleavage of rSERA6 by rPfSUB1. This might imply that SERA6 is not expected to be still processed in the integrated mutant lines. However, it cannot be excluded that other endogenous proteases cleave SERA6 in case of an incomplete PfSUB1 proteolysis to ensure correct SERA6 processing. Again this could be tested with cloned lines of the cleavage site mutant parasites. Towards this SERA6 could be purified from the mutant parasite clones and cleaved with rPfSUB1 in vitro in order to investigate the processing pattern by comparison to wild-type SERA6. Shedding light on SERA6 protease activity will provide insights into the importance of PfSUB1 processing.

4.3.4 Future tools to dissect SERA6 processing in the parasite are required

SERA5 processing in the parasite was first identified by the accumulation of the 50 kDa fragment in culture supernatants. These first results were followed by pulse-chase experiments combined with a time course that revealed the pattern of SERA5 processing in parasites (Delplace *et al.*, 1985;Delplace *et al.*, 1987;Debrabant *et al.*, 1992). SERA5 is one of the most abundant proteins of the parasite and tracking of various fragments was possible due to the fact that all fragments were readily visible on Coomassie Blue stained SDS-polyacrylamide gels following affinity purification. In addition the characterisation of SERA5 processing was facilitated by highly specific antibodies which improved immunoprecipitation after pulse-chase experiments. In this work, most approaches that led to the characterisation of SERA5 processing in the parasite were tried but were unsuccessful for SERA6. Additionally antibodies produced in this work were not suitable for immunoprecipitation. Pulse-chase experiments were conducted during this work but many proteins and not exclusively SERA6 were

immunoprecipitated (data not shown). Additional approaches to epitope-tag the C-terminus of *sera6* in the parasite were unsuccessful (Michael Shea, unpublished).

To entirely characterise SERA6 processing in the parasite two things have to be achieved: firstly, improvement of the SERA6 antibodies to enable future SERA6 immunoprecipitation from the parasite; secondly the detection of parasite SERA6 by tagging the genomic locus. Recent work by Stallmach *et al*. successfully replaced the 6 kDa fragment of SERA5, between SERA5st2 and SERA5stX, with a tandem affinity purification (TAP)-tag in the parasite (Robert Stallmach, unpublished). As the replacement of the 6 kDa fragment appeared to be tolerated by the parasite this might be a future approach for tagging endogenous SERA6 and following its processing. However, replacing the 6 kDa fragment might have an impact on SERA6 function as a putative protease and has to be considered with caution. Additionally, antibodies raised against the 6 kDa fragment may also improve the understanding of SERA6 processing in the parasite.

In summary: the improvement of antibody specificity, tagging the endogenous SERA6 and the production of new antibodies specific for the 6 kDa SERA6 fragment could potentially provide new tools to dissect SERA6 processing events in the parasite.

4.3.5 PbSERA3 is a substrate for PfSUB1

As the recombinant expression of SERA6 remained challenging PbSERA3 was hoped to be applicable as a SERA6 model in order to determine protease activity of the "Cys-type" SERAs. PbSERA3 was recently shown to be located in the PV of liver stage schizonts. It was also shown to be expressed in blood stages of the rodent malaria parasite (Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009). To ensure that PbSERA3 was in fact not only the SERA6 orthologue but also displayed similar processing patterns, the recombinant protein was subjected to rPfSUB1 processing assays. Recombinant PbSERA3 WT and the corresponding C639A mutant were found to be both processed by rPfSUB1 in a similar manner to SERA6. Intriguingly they were then further converted to a smaller form (~48 kDa). This form still contained the putative catalytic region, as it was detected by antibodies against the central PbSERA3 domain. N-terminal sequencing revealed that this was due to processing by PfSUB1 at a previously unpredicted site. Cleavage at this site was probably due to the presence of a small residue in P2 position and some amino acids with acidic and charged tendency in the ′P-positions probably allowed processing of this unusual site (Koussis, 2009;Koussis *et al.*, 2009;Silmon de Monerri, 2010;Silmon de Monerri *et al.*, 2011). This step was not observed in the rSERA6. It could therefore be argued that these results make PbSERA3 an inappropriate SERA6 model. However, these findings turned later out to be very informative and will be thoroughly analysed and discussed in

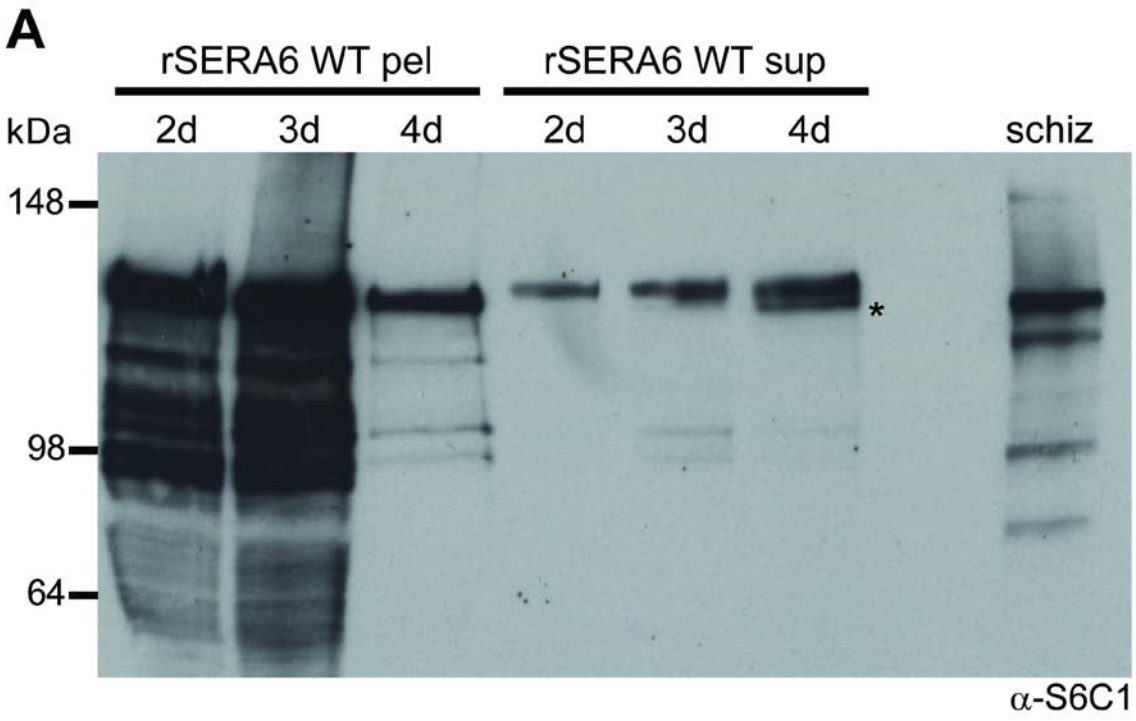
chapter 5. It is unknown whether *P. berghei* SUB1 processes PbSERA3 in an identical manner to PfSUB1, but fluorogenic peptide substrates based on PbSERA3st1 and PbSERA3st2 were processed by PbSUB1 *in vitro* (Catherine Suarez, unpublished).

As SERA6 and PbSERA3 undergo PfSUB1 processing in a similar manner, it was decided that rPbSERA3 will be applied as a model for SERA6 in future work and in the next part of this study.

Table 14: Modifications introduced to SERA6 processing sites

Transfection constructs	***Subsite***	***Modification***
pHH-SERA6chim	SERA6st1: VKAQ↓DDFN SERA6st2: VHGQ↓SNES SERA6stX: VYFL↓RHNP	Wild-type
pHH-SERA6chim_mutS1	SERA6st1: VK**LL**DDFN SERA6st2: VHGQ↓SNES SERA6st2: VYFL↓RHNP	site 1 P_{2-1}→ LL
pHH-SERA6chim∆S2	SERA6st1: VKAQ↓DDFN SERA6st2: -SNES SERA6stX: VYFL↓RHNP	site 2 deletion→ deletion of P_{4-1}
pHH-SERA6chim_mutS1∆S2	SERA6st1: VK**LL**DDFN SERA6st2: -SNES SERA6stX: VYFL↓RHNP	site 1 P2-1→ LL, site 2 deletion→ deletion of P4-1
pHH-SERA6chim_mutSx	SERA6st1: VKAQ↓DDFN SERA6st2: VHGQ↓SNES SERA6stX: V**AAA**RHNP	site X P3-1→ AAA
pHH-SERA6chim_mutS2	SERA6st1: VKAQ↓DDFN SERA6st2: VH**LL**SNES SERA6stX: VYFL↓RHNP	site 2 P2-1→ LL
pHH-SERA6chim_mutS2Sx	SERA6st1: VKAQ↓DDFN SERA6st2: VH**LL**SNES SERA6stX: V**AAA**RHNP	site 2 P2-1→ LL, site X P3-1→ AAA
pHH-SERA6chim_mutS1S2	SERA6st1: VK**LL**DDFN SERA6st2: VH**LL**SNES SERA6stX: VYFL↓RHNP	site 1 P2-1→ LL, site 2 P2-1→ LL
pHH-SERA6chim_mutS1S2Sx	SERA6st1: VK**LL**DDFN SERA6st2: VH**LL**SNES SERA6stX: V**AAA**RHNP	site 1 P2-1→ LL, site 2 P2-1→ LL site X P3-1→ AAA
pHH-SERA6chim_mutS1Sx	SERA6st1: VK**LL**DDFN SERA6st2: VHGQ↓SNES SERA6stX: V**AAA**↓RHNP	site 1 P2-1→ LL, site X P3-1→ AAA

Figure 30: Recombinant SERA6 is secreted into COS-7 cell culture supernatant
Recombinant SERA6 WT was transiently expressed from two to four days in COS-7 in culture medium not containing any fetal calf serum but in the presence of tunicamycin. The culture supernatant (sup) was harvested, pelleted and several fold concentrated. The remaining attached COS-7 cells were washed in PBS and harvested by scraping the cells off the culture flask. The cells were then collected and lysed in SDS sample buffer. Sup and COS-7 cells (pel) were separated by SDS-PAGE and analysed by Western blot using rabbit α-S6C1, total schizont extracts served as positive control (schiz). The majority of rSERA6 WT was expressed intracellular (A, rSERA6 WT pel). Optimal levels of secreted protein expression were obtained after four days (A, rSERA6 WT sup). The appearance of two bands in rSERA6 WT supernatants may have been due to glycosylation, indicating an insufficient concentration of tunicamycin. The lower band which appeared in rSERA6 WT sup samples after four days (A, asterisk) comigrates with full-length SERA6 from the schiz control. This band might display unglycosylated SERA6. Next, rSERA6 WT and rSERA6 C644A were expressed in the presence of a 2 fold increased tunicamycin concentration. The culture supernatants were harvested after four days, concentrated 16-fold and analysed by Western blot using rabbit α-S6C1 (B). Only one band appeared this time suggesting that tunicamycin concentration was sufficient.


A
rSERA6 WT pel
rSERA6 WT sup
kDa
2d
3d
4d
2d
3d
4d
schiz
148
98
64
*
α-S6C1

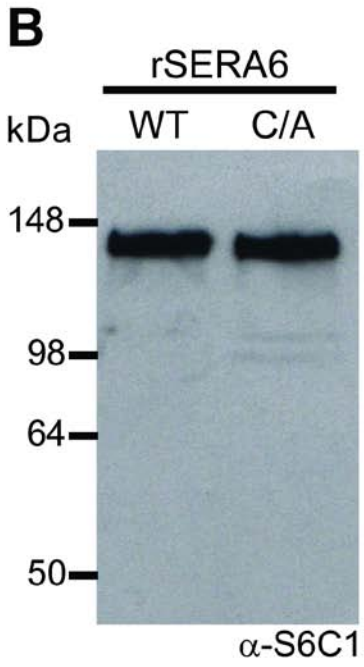

B
rSERA6
kDa
WT
C/A
148
98
64
50
α-S6C1

Figure 31: Recombinant and soluble SERA6 is processed by PfSUB1 at at least two sites

After the expression and secretion in COS-7 cell culture supernatant rSERA6 WT and rSERA6 C644A were processed with rPfSUB1. The rSERA6 containing COS-7 cell culture supernatants were concentrated and resuspended in PfSUB1 activity buffer and incubated in the presence of rPfSUB1. As controls both rSERA6 were incubated at 37°C alone or in presence of rPfSUB1 and the rPD. Samples were taken after 0, 10, 30 and 120 min, separated by SDS-PAGE followed by Western blot analysis with rabbit α-S6C1. Full-length SERA6 (~135 kDa) was gradually processed to a fragment at ~75 kDa (c) over two intermediates at ~125 kDa (a) and ~100 kDa (b). Recombinant SERA6 was stable by itself and mainly stable in the presence of the rPD. Slight processing occurred due to reduced inhibitory activity of the rPD. The central SERA6 antibody cross-reacted with the rPD which derived from the fact that the antibody was raised against a 6xHis-tag recombinant SERA6 and the rPD also carries a 6xHis-tag (asterisk). rPD: recombinant PfSUB1 prodomain

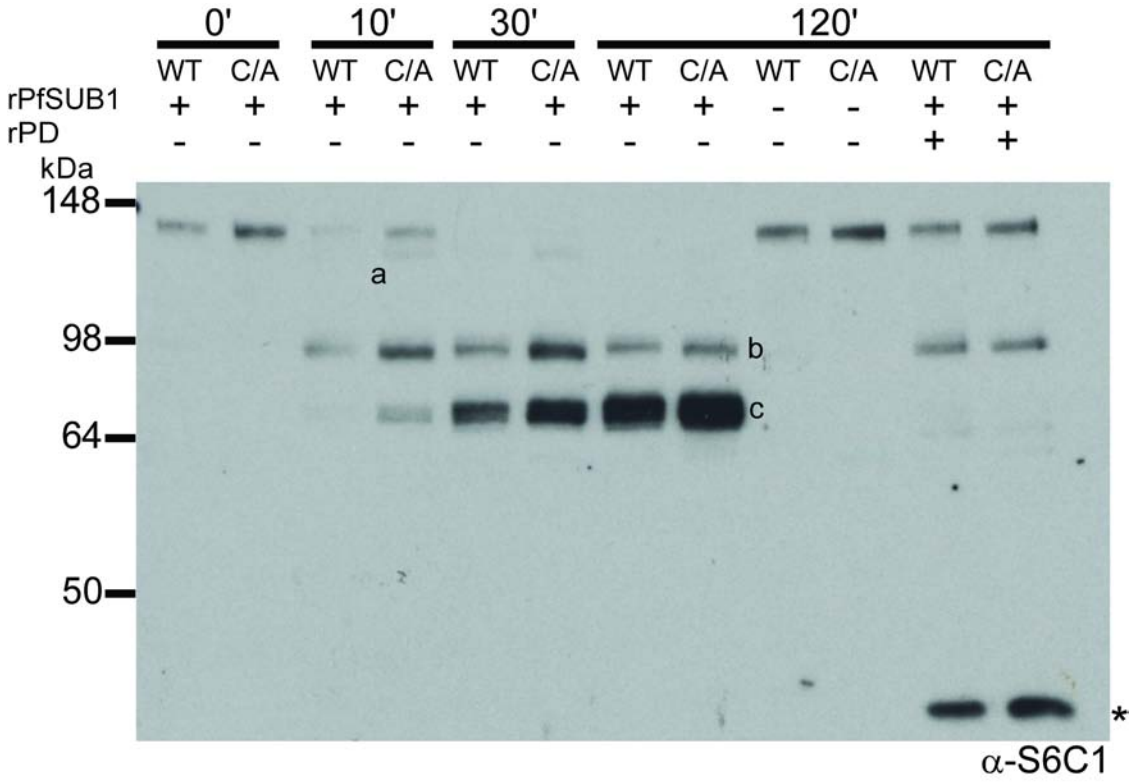

0'
10'
30'
120'
WT C/A WT C/A WT C/A WT C/A WT C/A WT C/A
rPfSUB1 + + + + + + + + - - + +
rPD - - - - - - - - - - + +
kDa
148
98
64
50
a
b
c
*
α-S6C1

Figure 32: Parasite-derived SERA6 can be purified from hypotonically lysed schizont extracts

In order to analyse SERA6 processing the protein was purified from hypotonically lysed parasites. Schizonts were hypotonically lysed and the extracts fractionated by size-exclusion chromatography. Recovered samples were then analysed by Western blot for the presence of SERA6 (A) and SERA5 (B) by using rabbit α-S6C1 or full-length SERA5 (mAb 24C6.1F1). The majority of SERA6 was detected in fractions 33 and 34 whereas the majority of SERA5 appeared in fraction 35. Fractions 33 and 34 were then applied in anion-exchange chromatography for further purification.

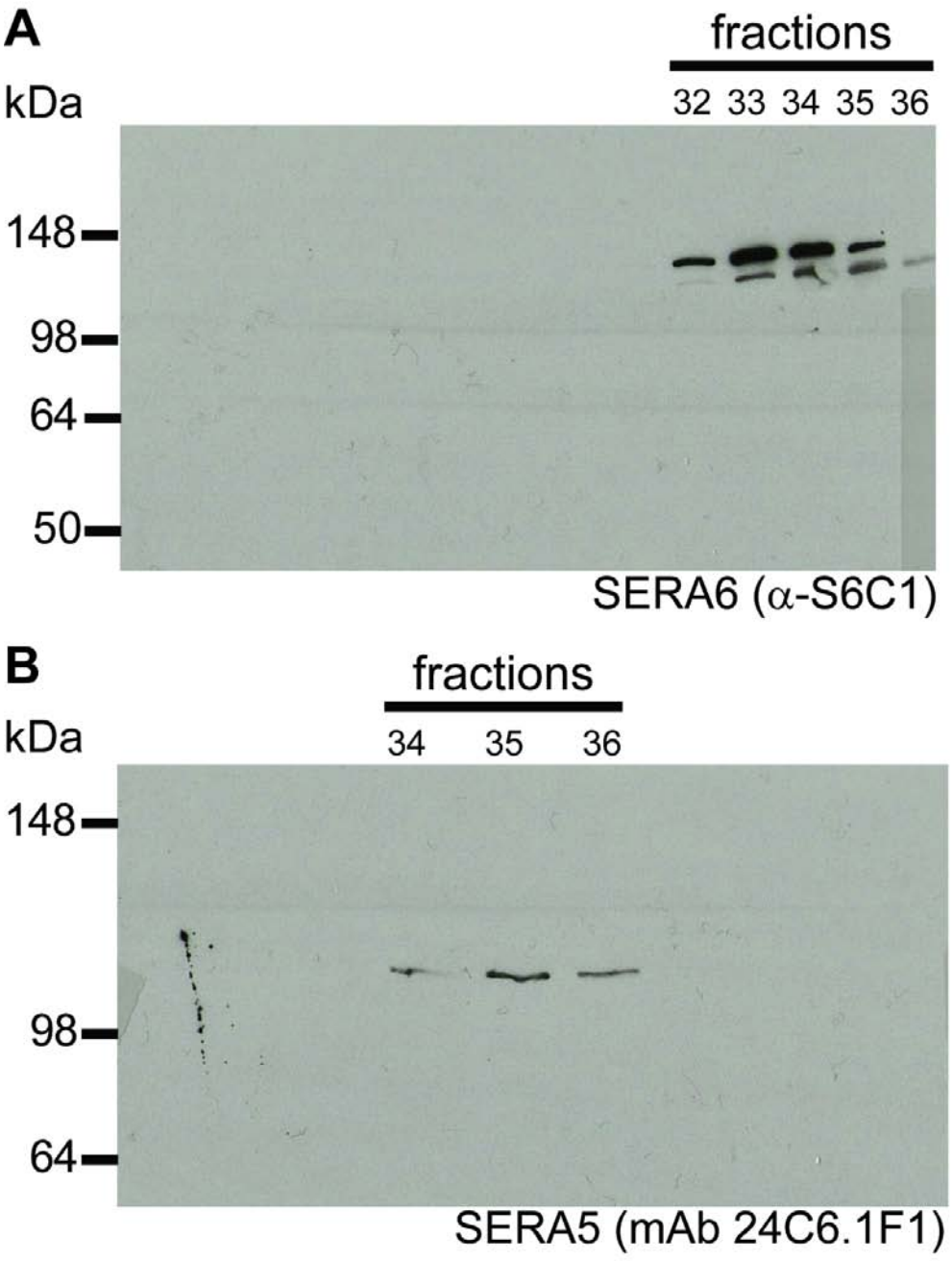
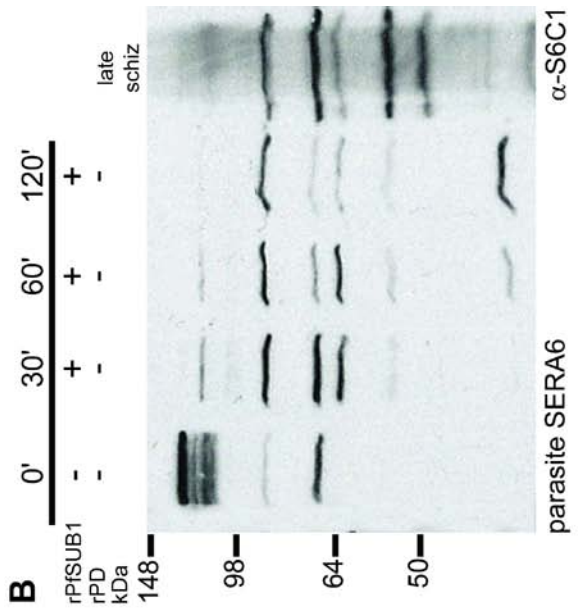
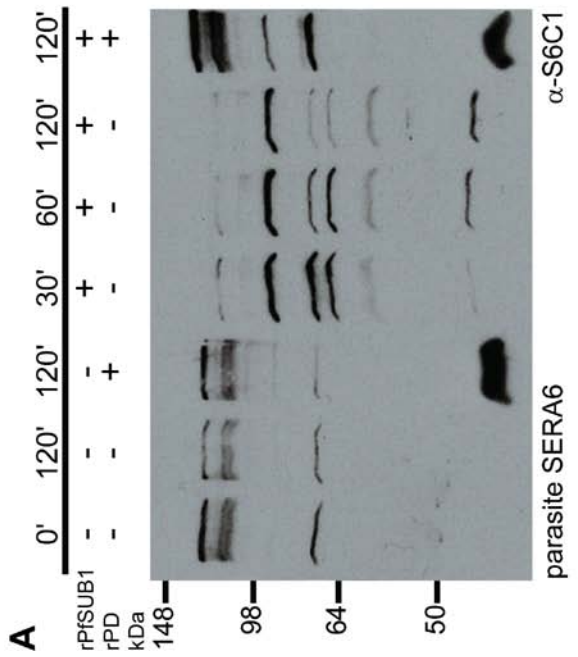

A
fractions
kDa
32 33 34 35 36
148
98
64
50
SERA6 (α-S6C1)
B
fractions
kDa
34 35 36
148
98
64
SERA5 (mAb 24C6.1F1)

Figure 33: Parasite-derived SERA6 is processed by PfSUB1 at at least two sites
To investigate whether parasite-derived SERA6 undergoes processing by rPfSUB1, parasite-derived SERA6 was incubated at 37°C with and without rPfSUB1, with or without rPD and combined with rPfSUB1 and rPD. Samples were taken at 0, 10, and 120 min and analysed by Western blot using rabbit α-S6C1 antibodies (A). SERA6 is cleaved several times, from a ~135 kDa product, via ~125, ~100, and ~75 kDa products to even smaller fragments. Minimal processing was observed in samples containing rPfSUB1 and the rPD. This was caused by the addition of rPfSUB1 moments before adding rPD. In a next step it was analysed whether the processing of the parasite-derived SERA6 was analogous to processing of SERA6 in the parasite. For that the samples from the rPfSUB1 processing assay were comigrated next to mature schizont extracts. The cleavage products migrate at the same size as the different SERA6 species found in very late schizonts (B, late schiz, lysed in SDS sample buffer), suggesting that *in vitro* cleavage by rPfSUB1 replicates the processing observed in the parasite. Additional und unpredicted processing of SERA6 was observed. To identify whether the COS-7 cell derived rSERA6 WT processed with rPfSUB1 also comigrates with SERA6 fragments from schizont extracts, rSERA6 WT was incubated at 37°C with rPfSUB1 and samples were taken at 0, 10, 60 and 180 min and analysed by Western blot using rabbit α-S6C1 (C). Samples were analysed by comigration with schizont extracts. Four major bands were observed ~135 kDa (C, a), ~125 kDa (C, b), ~100 kDa (C, c) and ~75 kDa (C, d). These results also indicate that the additional processing in parasite-derived SERA6 might be due to other proteases that were possibly copurified with SERA6. rPD: recombinant PfSUB1 prodomain


B
rPfSUB1
rPD
kDa
0'
30'
60'
120'
late schiz
148
98
64
50
parasite SERA6
α-S6C1


A
rPfSUB1
rPD
kDa
0'
120'
120'
30'
60'
120'
120'
148
98
64
50
parasite SERA6
α-S6C1

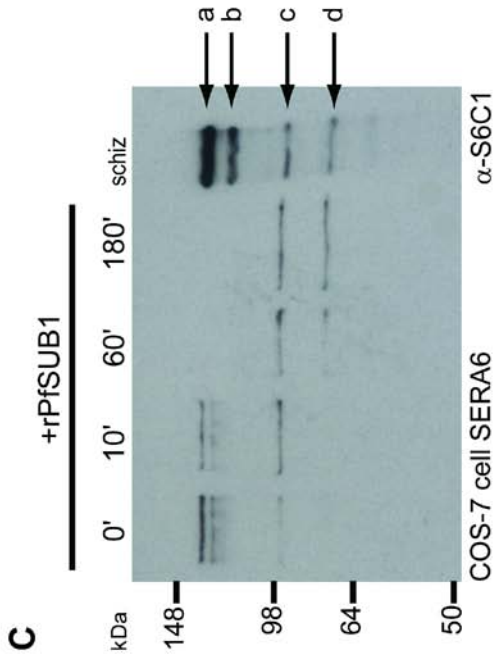

C
+rPfSUB1
kDa
0'
10'
60'
180'
schiz
a
b
c
d
148
98
64
50
COS-7 cell SERA6
α-S6C1

Figure 34: Strategy for N-terminal sequencing of SERA6 processing sites

Schematic of SERA6 showing the wild-type form (rS6-FL) and the mutations introduced to generate mutants SERA6mutS1 (rS6-FLmutS1), SERA6mutS2 (rS6-FLmutS2), and SERA6mutS1S2 (rS6-FLmutS1S2) for recombinant expression (A). These mutants were designed to block PfSUB1 processing. As attempts to purify rSERA6 remained unsuccessful a new strategy was developed to obtain pure rSERA6 (B). Various recombinant SERA6 mutants were expressed in *E. coli* as IB*,* IB were then lysed with 8 M urea and samples were buffer exchanged to PBS. The samples were then separated by SDS-PAGE and the bands corresponding to full-length SERA6 were excised and the crushed gel pieces incubated with rPfSUB1. The digestion products were again separated by SDS-PAGE, transferred to PVDF membrane, and the cleavage sites were identified via N-terminal sequencing. IB: *E. coli* inclusion bodies

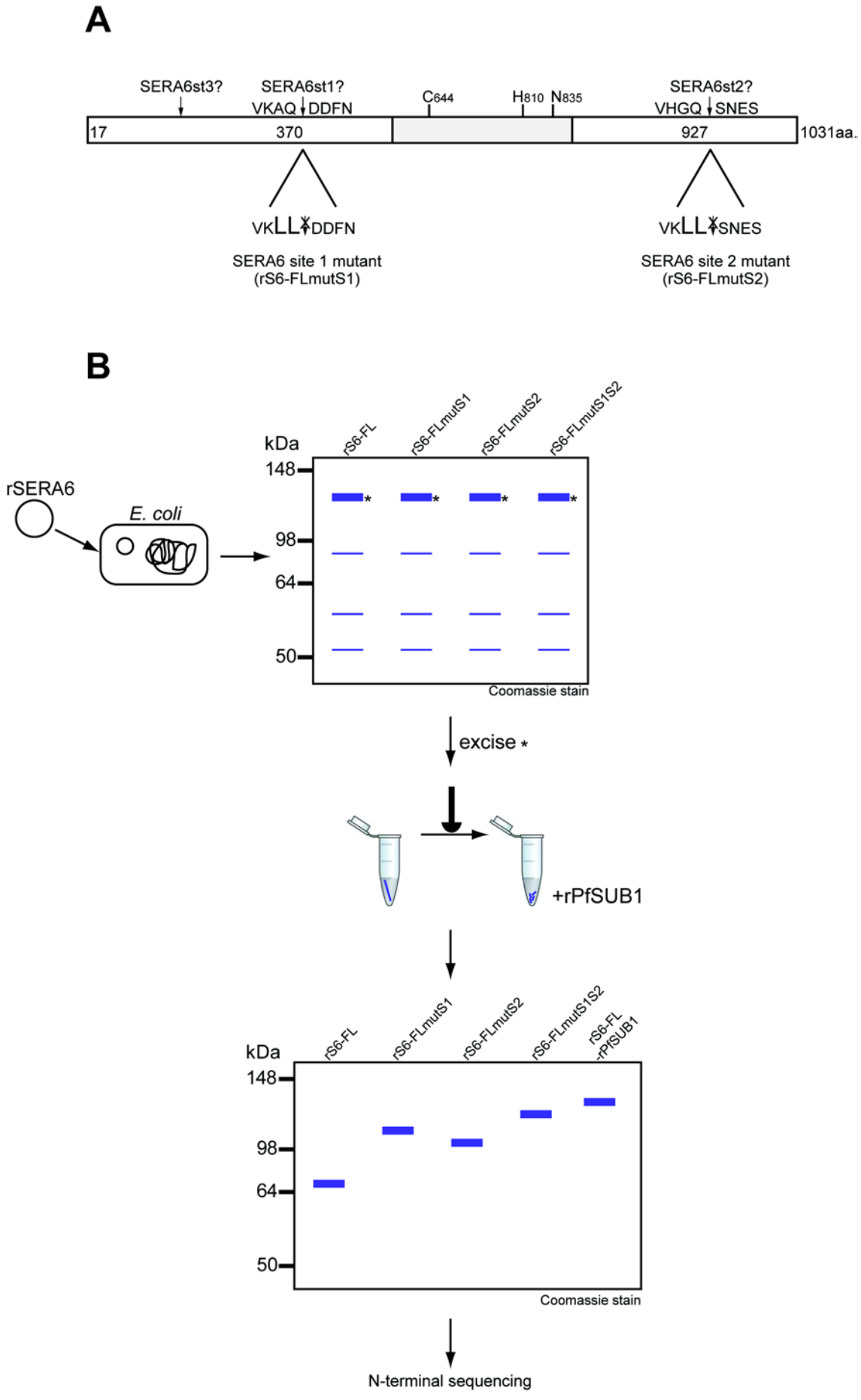
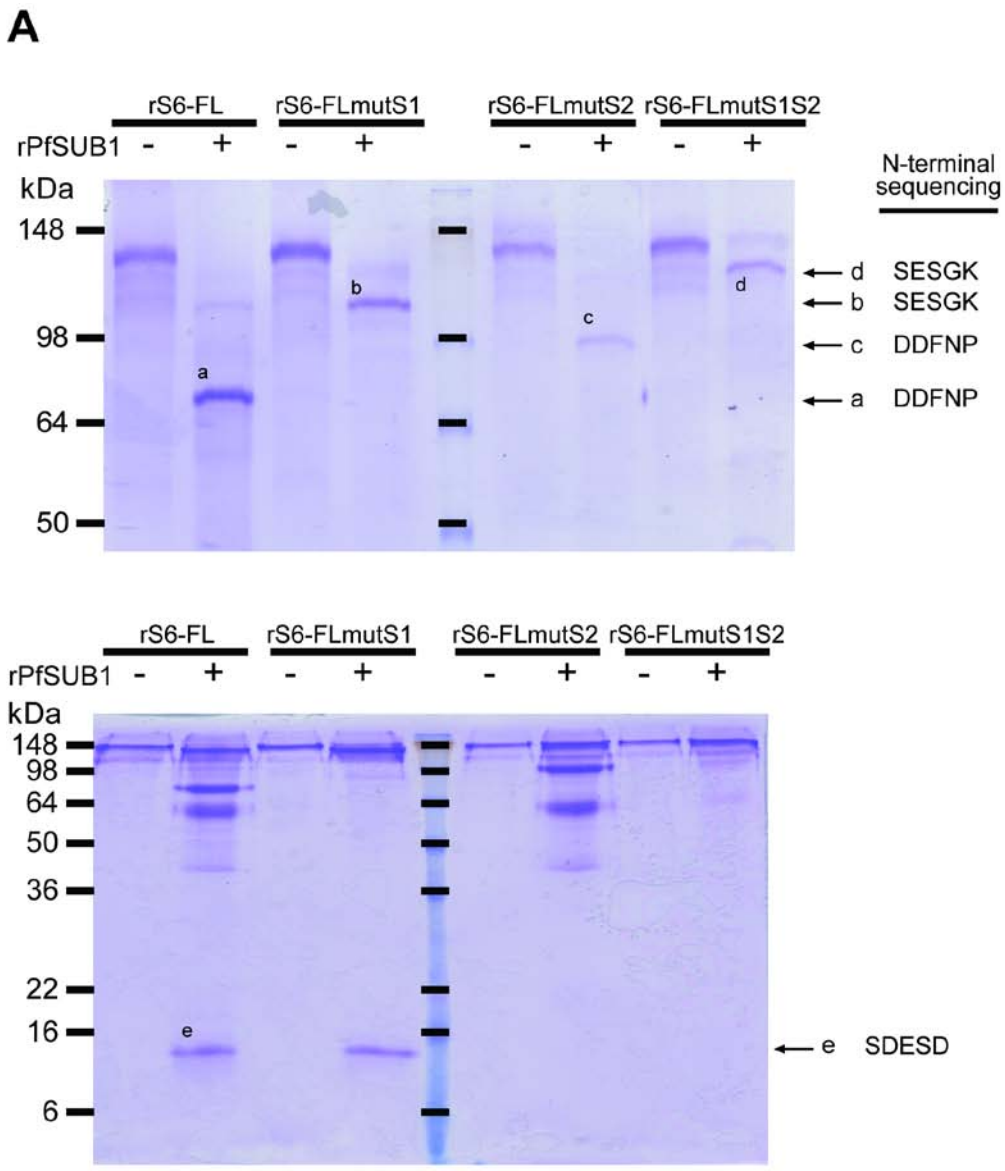

A
SERA6st3?
SERA6st1?
VKAQ↓DDFN
C644
H810 N835
SERA6st2?
VHGQ↓SNES
17
370
927
1031aa.
VKLL↓DDFN
VKLL↓SNES
SERA6 site 1 mutant
(rS6-FLmutS1)
SERA6 site 2 mutant
(rS6-FLmutS2)
B
rSERA6
E. coli
kDa
148
98
64
50
rS6-FL
rS6-FLmutS1
rS6-FLmutS2
rS6-FLmutS1S2
Coomassie stain
excise *
+rPfSUB1
rS6-FL
-rPfSUB1
N-terminal sequencing

Figure 35: Recombinant SERA6 is cleaved in three distinct sites by PfSUB1 *in vitro*

Recombinant full-length proteins SERA6 (rS6-FL), SERA6mutS1 (rS6-FLmutS1), SERA6mutS2 (rS6-FLmutS2), and SERA6mutS1S2 (rS6-FLmutS1S2) were incubated with or without rPfSUB1 and products visualised on a Coomassie Blue stained SDS-PAGE gel (A). Top panel: Full-length SERA6 is processed to a ~75 kDa form (a). SERA6mutS1 is processed to a ~110 kDa form (b). SERA6mutS2 is processed to a ~100 kDa form (c). SERA6mutS1S2 is unexpectedly processed to a ~125 kDa form (d). Bands a, b, c and d were N-terminally sequenced and the obtained sequence is indicated to the right of the gel. Bottom panel: a ~12 kDa product was generated by rPfSUB1 cleavage of recombinant SERA6 and SERA6mutS1 (e) and shown to be the C-terminal fragment by N-terminal sequencing. N-terminal sequencing confirmed two previously predicted processing sites and identified a third previously unpredicted processing site (B).


A
rS6-FL
rS6-FLmutS1
rS6-FLmutS2
rS6-FLmutS1S2
rPfSUB1 - + - + - + - +
kDa
148
98
64
50
N-terminal sequencing
a
b
c
d
d SESGK
b SESGK
c DDFNP
a DDFNP
rS6-FL
rS6-FLmutS1
rS6-FLmutS2
rS6-FLmutS1S2
rPfSUB1 - + - + - + - +
kDa
148
98
64
50
36
22
16
6
e
e SDESD

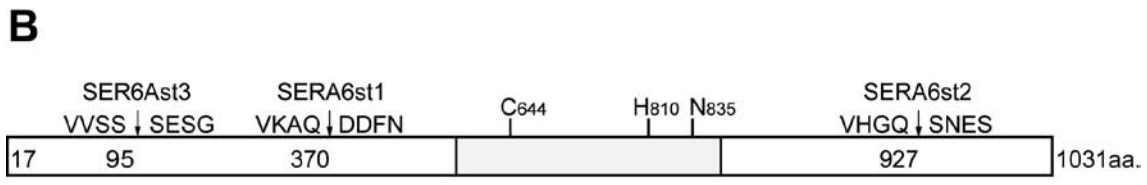

B
SER6Ast3
VVSS ↓ SESG
SERA6st1
VKAQ ↓ DDFN
C644
H810 N835
SERA6st2
VHGQ ↓ SNES
17
95
370
927
1031aa.

Figure 36: Processing of SERA6 by PfSUB1 in at least one site is indispensable in the parasite

Schematic of transfection strategy for single homologous recombination applied in this work (A). The strategy was based on a chimera between genomic *sera6* and recodonised *sera6* which facilitated single-homologous recombination in the targeted *sera6* locus. The recodonised portion of each construct carried the cleavage site mutations. Two primer pairs were applied for identification of positive integrants by diagnostic PCR with the primer binding sites labelled in green (A). Primer pair 5′UTRb and 9 were used for the confirmation of the presence of genomic DNA. Primer pair 5′UTRb and 37 were applied to detect any integration (~1.8 kbp) as primer 37 exclusively binds to the recodonised *sera6* sequence (A).

Results of the diagnostic PCR of SERA6 chimera wild-type control (pHH-SERA6chim) and parasite lines showing presence of homologous recombination events are shown in B. In a first step SERA6st2 deletion of P4-P1 (pHH-SERA6chim∆S2), SERA6st2 substitution of P2-P1 with Ala-Ala (pHH-SERA6chim_mutS2) and SERA6st2 substitution in combination with SERA6stX substitution of P3-P1 with Ala-Ala-Ala (pHH-SERA6chim_mutS2Sx) were identified as positive for integration by diagnostic PCR (B, left hand panel). At the same time the presence of genomic DNA was confirmed (B, left hand panel, asterisk). A pHH-AMA1 transfected and cloned line served as negative control. A doublet was always observed, similar to SERA5 (Robert Stallmach, unpublished). The PCR positive lines were then analysed by Southern blot using the DIG method (C). Bands at ~3.3 kbp are expected for if the integration of the SERA6 chimera was successful. All by PCR identified integration events were confirmed by Southern blot (C). The SERA6 chimera transfection plasmid served as control (pHHSERA6chimplasmid). The transfected lines were not cloned therefore the plasmid and the non-integrated *sera6* loci were still detected by Southern blot (C, both ~9.5 kbp, asterisk).

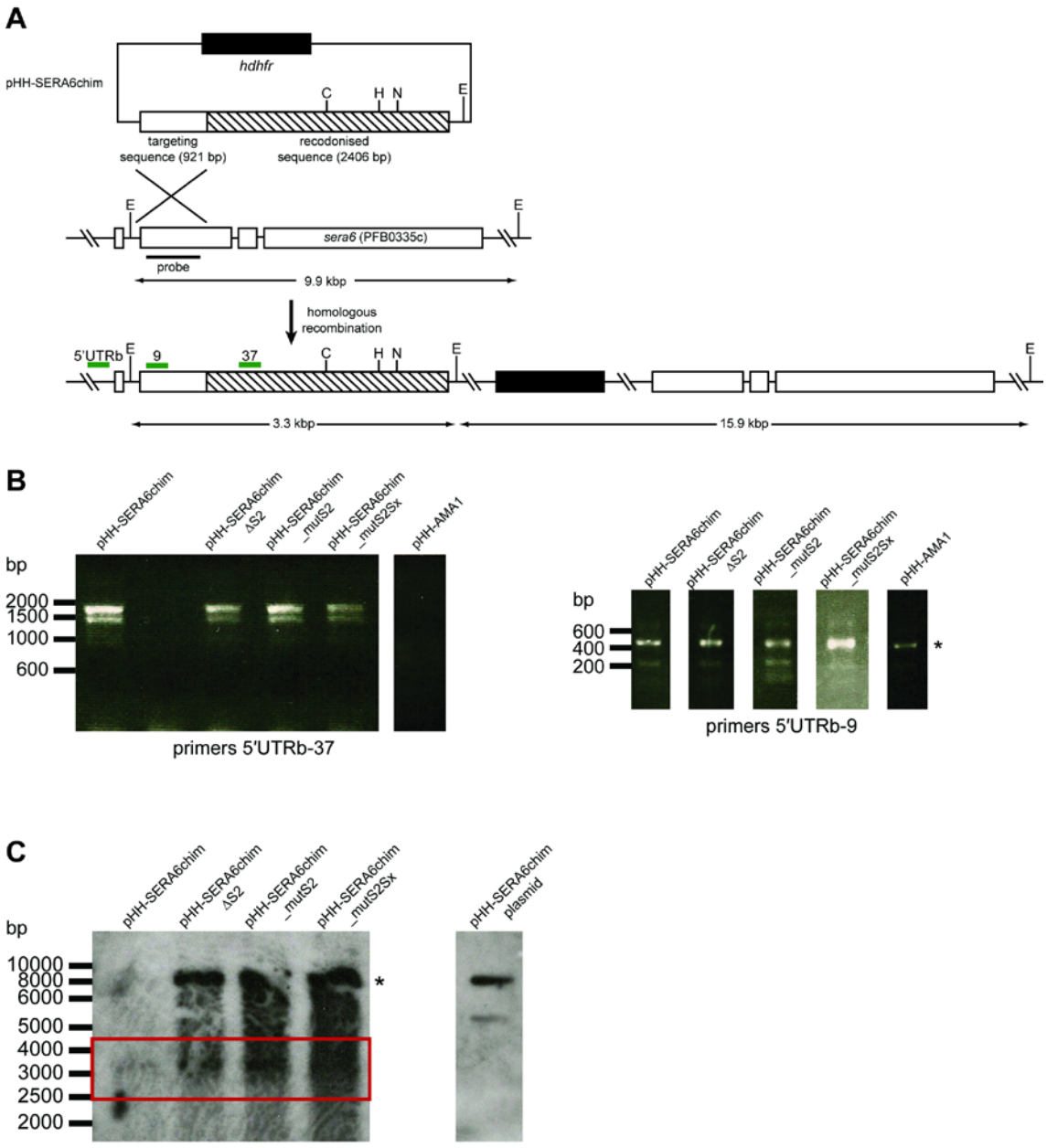

A
pHH-SERA6chim
hdhfr
C
H N
E
targeting sequence (921 bp)
recodonised sequence (2406 bp)
E
E
sera6 (PFB0335c)
probe
9.9 kbp
homologous recombination
5'UTRb
E
9
37
C
H N
E
E
3.3 kbp
15.9 kbp
B
pHH-SERA6chim
pHH-SERA6chim ΔS2
pHH-SERA6chim _mutS2
pHH-SERA6chim _mutS2Sx
pHH-AMA1
bp
2000
1500
1000
600
primers 5'UTRb-37
pHH-SERA6chim
pHH-SERA6chim ΔS2
pHH-SERA6chim _mutS2
pHH-SERA6chim _mutS2Sx
pHH-AMA1
bp
600
400
200
*
primers 5'UTRb-9
C
pHH-SERA6chim
pHH-SERA6chim ΔS2
pHH-SERA6chim _mutS2
pHH-SERA6chim _mutS2Sx
bp
10000
8000
6000
5000
4000
3000
2500
2000
*
pHH-SERA6chim plasmid

Figure 37: Recombinant PbSERA3 is processed by recombinant PfSUB1 in a SERA6 similar manner

Recombinant PbSERA3 WT and rPbSERA3 C639A were expressed in insect cells. Culture supernatants were harvested after three days, pelleted and filtered. Supernatants were then purified by anion-exchange chromatography followed by size-exclusion chromatography. Recovered samples were then analysed by Western blot using α-PbS3C1. Both recombinant proteins were eluted in fractions 31 to 33 (A). The fractions were then pooled separately for each protein, buffer exchanged into PfSUB1 activity buffer and subjected to rPfSUB1 processing assays. The rPbSERA3 WT and rPbSERA3 C639A were incubated at 37°C with or without rPfSUB1, with or without rPD and combined with rPfSUB1 and rPD. Samples were taken at 0 and 120 min and analysed by Western blot using α-PbS3C1. Recombinant PbSERA3 was processed by rPfSUB1 from ~135 kDa (B, a) over ~100 kDa (B, b) to ~75 kDa (B, c). Additional processing was observed for rPbSERA3 WT to ~48 kDa (B, d) and ~42 kDa (B, e). For rPbSERA3 C639A only the conversion to the ~48 kDa (B, d) was observed. The conversion of both rPbSERA3 to a ~48 kDa fragment was unpredicted as an event mediated by rPfSUB1. Recombinant PbSERA3 underwent unexpected processing in samples incubated with rPfSUB1 and rPD. Fluorogenic substrate peptide assays revealed that the rPD was no longer functional and was replaced with a fresh stock rPD.

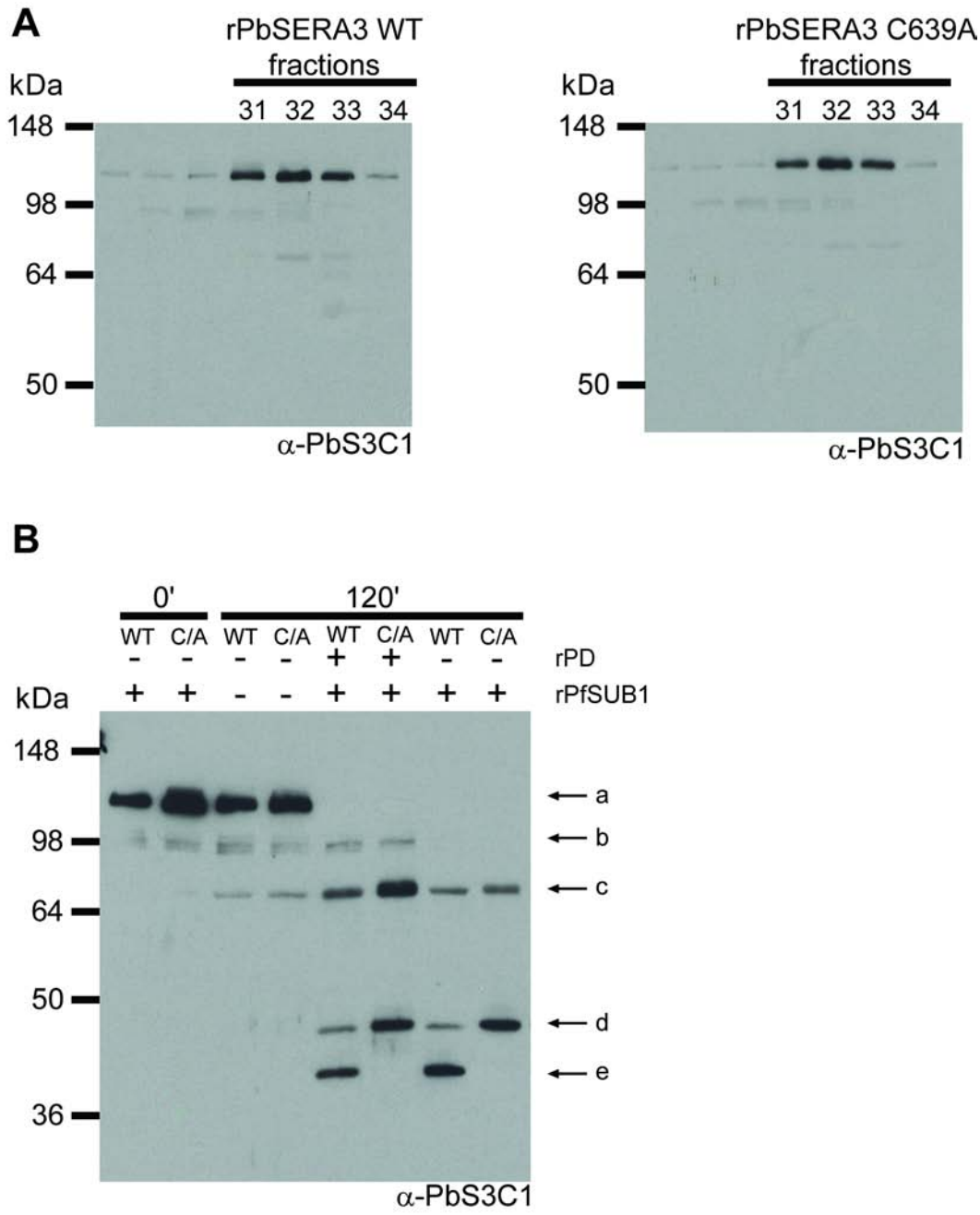
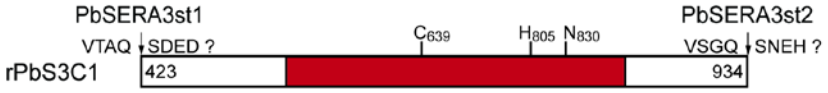

A
rPbSERA3 WT
fractions
kDa
31 32 33 34
148
98
64
50
α-PbS3C1
rPbSERA3 C639A
fractions
kDa
31 32 33 34
148
98
64
50
α-PbS3C1
B
0'
120'
WT C/A WT C/A WT C/A WT C/A
- - - - + + - - rPD
+ + - - + + + + rPfSUB1
kDa
148
98
64
50
36
a
b
c
d
e
α-PbS3C1

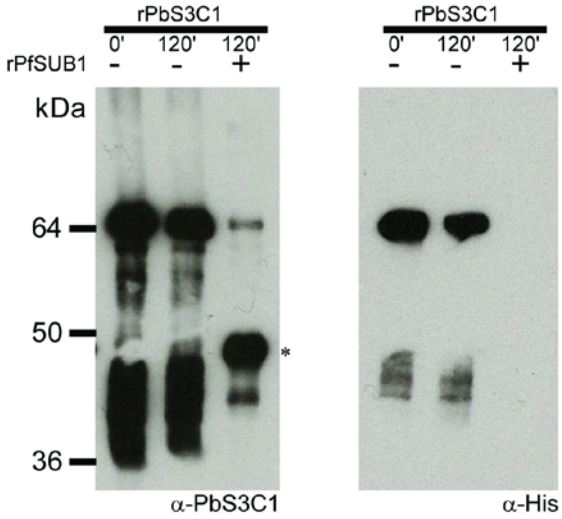
Figure 38: PbSERA3 is processed by PfSUB1 in an unpredicted site

The appearance of a 48 kDa rPbSERA3 fragment after processing with rPfSUB1 raised the question of whether PbSERA3 encodes for an unpredicted PfSUB1 processing site. An rPbSERA3 fragment, between the predicted PbSERA3st1 and PbSERA3st2, was expressed with an N-terminal 6xHis- and S-tag (rPbS3C1, A). The recombinant protein was expressed as IB. The IB were solubilised with 8 M urea and rPbS3C1 was purified using Nickel-Histidine chelation. The recombinant protein was then subjected to rPfSUB1 processing assays, samples were taken at 0 and 120 min and analysed by Western blot using α-PbS3C1. As expected the ~75 kDa fragment was processed to a fragment of 48 kDa (B, asterisk). To investigate whether the cleavage occurred on the N- or C-terminus the same nitrocellulose membrane was probed with a α-His antibody (the rPbS3C1 carried an N-terminal 6xHis-tag). The signal disappeared after processing by rPfSUB1 implying that cleavage occurs at the N-terminus. The same samples were transferred to PVDF membrane and the processing site was N-terminally sequenced. Recombinant PfSUB1 processes rPbSERA3 also at an unfavourable AIGQ↓NEEP motif (C).

A


PbSERA3st1
VTAQ ↓ SDED ?
C639
H805 N830
PbSERA3st2
VSGQ ↓ SNEH ?
rPbS3C1
423
934

B


rPbS3C1
0' 120' 120'
rPfSUB1 - - +
kDa
64
50
36
*
α-PbS3C1
rPbS3C1
0' 120' 120'
- - +
α-His

C

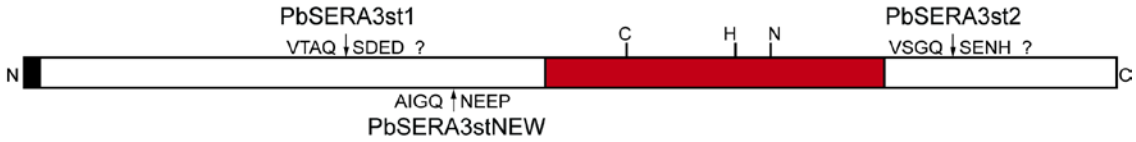
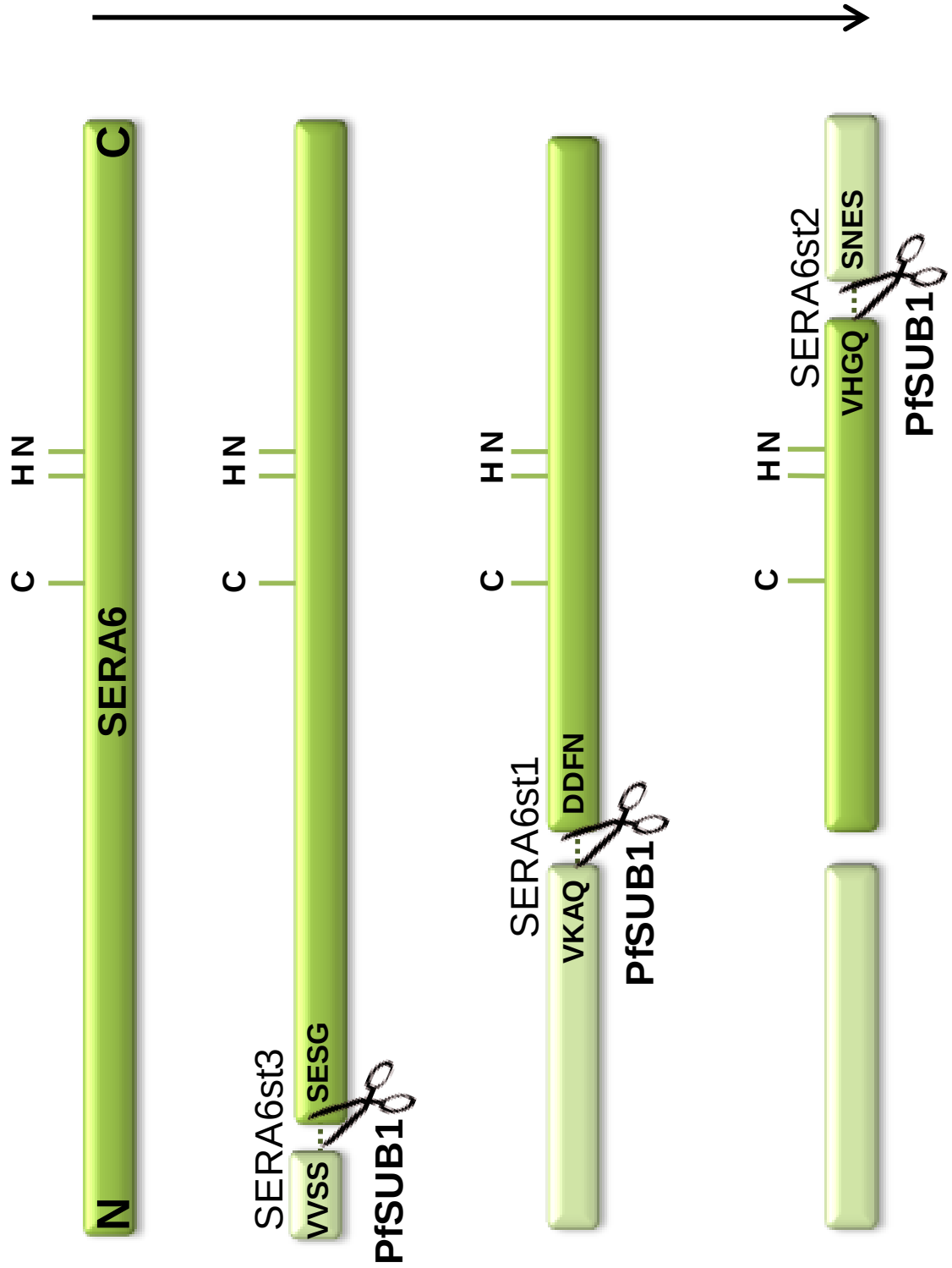

PbSERA3st1
VTAQ ↓ SDED ?
C
H
N
PbSERA3st2
VSGQ ↓ SENH ?
N
C
AIGQ ↑ NEEP
PbSERA3stNEW

Figure 39: Overview of SERA6 processing by PfSUB1

SERA6 is processed by PfSUB1 at three sites. The full-length SERA6 was processed in site SERA6st3 to a ~125 kDa fragment. This was suggested to be the first step in a proteolytic cascade as this fragment disappeared throughout the course of the processing assay. SERA6 undergoes further processing by PfSUB1 in SERA6st1 followed by processing in SERA6st2 releasing fragments of ~100 kDa and ~75 kDa respectively. Both of these fragments carry the central papain-like protease domain. Additional processing steps have not been analysed in this work.


N
SERA6
C
H N
C
H N
C
SERA6st3
VVSS
SESG
PfSUB1
H N
C
SERA6st1
VKAQ
DDFN
PfSUB1
H N
C
SERA6st2
VHGQ
SNES
PfSUB1

5 PbSERA3 and SERA6 are malarial proteases that require activation by PfSUB1

5.1 Introduction

As discussed in the Introduction, SERA6 carries a central papain-domain which putatively confers a cysteine protease activity. Previous work showing the indispensable nature of the *sera6* gene in *P. falciparum* cultures (Miller *et al.*, 2002b) raised the possibility that this cysteine protease activity may be required for the role of SERA6 and parasite survival. In this part of the work this putative protease activity of SERA6 was thoroughly investigated. First, it was attempted to precisely analyse the importance of the putative catalytic Cys residue for parasite survival, by attempting to modify this residue in the parasite. This approach was intended to address two main questions: firstly, whether the putative catalytic Cys is required for the correct function of SERA6 and secondly whether this function is required for parasite survival. It was next attempted to determine the proteolytic activity of SERA6. In a first step *E. coli* derived recombinant SERA6 was refolded and analysed for protease activity. Secondly it was attempted to purify SERA6 from parasite culture supernatants to investigate protease activity analysis. This was followed by protease activity studies with the parasite-derived SERA6 and *E. coli* cell lysates containing soluble rS6C1. Finally, the putative cysteine protease activity of recombinant PbSERA3, the *P. berghei* orthologue of SERA6, was examined.

5.2 Results

5.2.1 The putative catalytic Cys of SERA6 is indispensable for parasite survival

Many previous studies have shown that the replacement of a catalytic residue in cysteine proteases can ablate proteolytic activity without affecting the protease fold (Baker *et al.*, 1993;Pei *et al.*, 1996). This approach has been used to investigate the activity of a variety of enzymes and putative enzymes. In order to establish the importance of the putative catalytic Cys644 of SERA6, a substitution of this residue with an Ala644 residue in 3D7 *P. falciparum* parasites was attempted. Predictions were that this Cys residue confers a cysteine protease function of SERA6 and a replacement would not be tolerated by the parasites. As demonstrated above a chimera construct of genomic *sera6* and recodonised *sera6* can be used to successfully drive single cross-over homologous recombination into the *sera6* locus (see chapter 4). Accordingly a similar transfection strategy was applied to integrate the catalytic residue substitution into the endogenous *sera6* locus in 3D7 parasites. Two constructs were designed, one (pHH-SERA6chim) designed to reconstitute the wild-type SERA6 gene product as a control for integration, and the other designed to produce the SERA6 C644A substitution (pHH-SERA6chimC/A, see Figure 40A, for a construct overview). Parasites

were transfected on 11 independent occasions with pHH-SERA6chim and seven independent occasions with pHH-SERA6chimC/A and integrated parasite populations were selected by cycling on and off drug. A diagnostic PCR was designed with primers to bind to recodonised *sera6* and which were unlikely to anneal to genomic *sera6* making the identification of integrant lines more specific (for additional information see chapter 4). As shown in Figure 40B, integration was detected for the SERA6 WT but not SERA6 C644A as a band of the expected size (1.8 kbp). Southern blot analysis was then performed in an attempt to confirm the integration of the PCR positive parasite lines. As shown in Figure 40C one band at ~3.3 kbp was detected in digests of genomic DNA from the SERA6 WT line but not of the SERA6 C644A line, indicating integration of SERA6 WT but not SERA6 C644A. SERA6 C644A on the other hand only showed a band at ~9.5 kbp (see Figure 40C, asterisk). These parasites appeared to have maintained the transfection vector in episomal form and also contain the undisrupted *sera6* locus. The detection of bands other than the integrated *sera6* locus was due to the fact that although the transfected lines had undergone drug cycling, they had not been cloned. These results suggest that the replacement of Cys644 of SERA6 is not tolerated by the parasite. Therefore the findings suggest that this putative catalytic residue is required for the persistence of the parasite asexual life-cycle.

5.2.2 Attempted refolding of recombinant SERA6

Having established that the predicted catalytic Cys644 of SERA6 appeared indispensable for parasite survival *in vitro* it was next attempted to demonstrate the protease activity of SERA6 *in vitro*. As recombinant expression of SERA6 remained challenging the screen for protease activity was attempted with a variety of approaches. An overview of these approaches can be found in Table 15. Many eukaryotic proteases were shown to successfully gain proteolytic activity after they were recovered from IB and then correctly refolded (Cabrita *et al.*, 2004). A commonly used method for protein refolding is rapid dilution as described for many proteins (Buckle *et al.*, 2005;Chow *et al.*, 2006). Therefore it was expected that once rSERA6 was correctly refolded it would become proteolytically active. Initial attempts to do this with SERA6 used rSERA6 which was expressed in *E. coli* IB. In a first step it was tried to recover various rSERA6 fragments from IB followed by several refolding attempts. The refolded rSERA6 fragments were then used to screen for protease activity via zymography. Zymograms are carried out by separating the putative protease on a SDS-polyacrylamide gel copolymerised with the enzyme substrate. Once activated by buffer exchanging, to remove the SDS, the protease can cleave the substrate to smaller fragments which can readily diffuse out from the gel. After staining the gel with Coomassie Blue this results in a band of clearing at the position of protease activity.

To attempt to obtain correctly refolded rSERA6, IB containing either rS6-FL, rS6C1 or rS6C2 were first washed and then solubilised in 8 M urea. All protein refolding attempts were made directly from the solubilised material without any preceding purification steps. All three proteins were then rapidly diluted into a large variety of refolding buffers (see Table 10). The refolded samples were then analysed by separation on CN-PAGE since insoluble protein aggregates would be expected to not enter the CN-polyacrylamide gels. This was followed by Western blot analysis using a commercial α-His antibody (recognising the N-terminal 6xHis-tag). Figure 41A shows that bands appeared for rS6C1 and rS6C2 after the recombinant proteins were refolded in the buffers that were used in refolding approaches for the following proteins: falcipain 2, SERA5 and also buffer 1 to 3 from the Quick Refolding Kit (Pierce). A band for rS6-FL was not as clearly visible but a smear was detected by Western blot. The results suggested that at least a small proportion of the three rSERA6 fragments was refolded. In an additional approach to assess the state of refolding the same samples were analysed by adding DTT or not to the 2x SDS sample buffer. Upon the addition of DTT the intramolecular disulphide-bonds of refolded proteins would be disrupted causing a size shift on SDS-polyacrylamide gels between samples with and without DTT. As seen in Figure 41B, samples that were refolded in falcipain 2 refolding buffer did not reveal a size shift upon the addition of DTT. This suggested that neither rS6C1 nor rS6C2 were correctly refolded. However, the Western blot results indicated refolding of at least a small proportion of rS6C1 and rS6C2. Therefore zymography was used to screen these samples for protease activity. Both rS6C1 and rS6C2 were refolded in the falcipain 2 refolding buffer and the samples were separated on SDS-polyacrylamide gels copolymerised with gelatin. After the electrophoresis the SDS was replaced with TX-100 and each gel was incubated in a buffer designed to allow cysteine protease activity (Chow *et al.*, 2006). As a control, 1 μg papain was additionally run on each of the gels. No white bands, except in the papain sample, were visible in Figure 41C indicating that rS6C1 and rS6C2 did not process any of the substrates. The most likely explanation for this observation was that neither the refolded rS6C1 nor rS6C2 possess any proteolytic activity in this assay. A second explanation could be that the recombinant fragments were correctly refolded but the substrates or developing buffers used were unsuitable. Additional substrates (BSA and marvel milk powder) and developing buffers (a pH 7.7 buffer or falcipain 2 activity buffer) were therefore tried, but also failed to reveal any protease activity (data not shown).

In summary in this section it was attempted to characterise SERA6 cysteine protease activity by refolding protein fragments that were recombinantly expressed in *E. coli* IB. Although a variety of methods was applied, protease activity was not

observed, which might have been due to several reasons. Firstly the proteins were likely to be only partially refolded, hence were not in their native state which is required for the catalytic triad residues to be in the correct proximity for efficient proteolytic activity. Secondly, in case the proteins had been correctly refolded it was unclear whether any of the substrates used in the zymography approaches were suitable substrates for SERA6. As it was difficult to assess why the fragments were proteolytically inactive it was decided to next screen for SERA6 activity with the parasite-derived protein.

5.2.3 Parasite-derived SERA6 shows no detectable protease activity

After the unsuccessful activity screen of refolded rSERA6, the aim of this section was to investigate further into the putative cysteine protease activity of SERA6 by using parasite-derived SERA6.

5.2.3.1 A fragment of SERA6 isolated from parasite culture supernatants shows no detectable protease activity

It is known that upon the rupture of iRBCs a small 50 kDa SERA5 fragment, containing the central papain-like protease domain, can be found in culture supernatant (Delplace *et al.*, 1985;Delplace *et al.*, 1987). It had yet to be established whether a similar SERA6-derived central domain can also be detected in culture supernatant after egress. Therefore it was first decided to attempt to identify and purify a SERA6 central domain from parasite culture supernatants after egress which could then be used to screen for activity.

To investigate whether a SERA6-central domain indeed accumulated in the culture supernatant of ruptured parasites, culture supernatants were collected and fractionated by anion-exchange chromatography as described by (Silmon de Monerri *et al.*, 2011). Fractions were analysed by Western blot using rabbit α-S6C1. As shown in Figure 42A an approximately 42 kDa protein reactive with the antibody was eluted between fractions 26 and 31 (see Figure 42A, left hand panel). These fractions were pooled, concentrated 60-fold and further purified by size-exclusion chromatography. The reactive protein (henceforth referred to as SERA6-central) was recovered in two peak fractions (see Figure 42A, right hand panel). These were pooled and stored froze in aliquots.

The next set of experiments involved screening the SERA6-central for protease activity. One of the approaches used was the incubation of the SERA6 central domain with iRBC ghosts. Infected RBC ghosts are lysed and resealed RBCs that still harbour all the membranes derived from the parasite and the RBC with the integrated and peripheral associated membrane proteins. Therefore it was suggested that if the

SERA6 central domain possessed proteolytic activity one of the membrane proteins found in iRBC ghosts might be processed by the putative cysteine protease.

In a first step SERA6-central was tested for autocatalytic processing. This was attempted by concentrating the SERA6-central 5-fold and incubating it for 4 hours or overnight at 37°C with or without the addition of calcium (some cysteine proteases, such as the calpains, are known to be calcium dependant Ohno *et al.*, 1984). The samples were then analysed by SDS-PAGE and Coomassie Blue staining. As shown in Figure 42B, no signs of autocatalytic breakdown were observed in any of the samples. In a second attempt to look for protease activity in the SERA6-central preparations, the protein was incubated with iRBC ghosts that were obtained by lysing iRBCs in a hypotonic buffer. Following incubation of iRBC ghosts in the presence of SERA6-central at 37°C, samples taken after 0, 1, 2, 4, 5 hours were analysed by SDS-polyacrylamide gels and silver staining. As seen in Figure 43A no differences were observed between the start sample and samples after 5 hours, suggesting no protease activity against the ghost proteins or the processed proteins were too low in abundance to be visible.

In a final approach the activity of SERA6-central was analysed with BSA substrate zymograms. SERA6-central was separated in polyacrylamide gels containing BSA and the gels incubated in four distinct buffers: falcipain 2 activity buffer, papain-activity buffer, and a pH 8.2 buffer with or without the addition of 5 mM $CaCl_2$. 1 μg papain was used as positive control. As seen in Figure 43B no white bands were visible for SERA6-central. This indicated that either SERA6-central was inactive or the conditions were unsuitable for SERA6-central to process BSA.

In conclusion, several attempts to detect protease activity in the parasite derived SERA6-central preparations were unsuccessful. This could be due to several reasons. One possible explanation may be that the SERA6-central domain found in culture supernatants does not possess proteolytic activity or the assay conditions were unsuitable for SERA6-central to show activity.

5.2.3.2 Parasite-derived full-length SERA6 possesses no detectable protease activity

The next set of experiments focused in assessing parasite-derived full-length SERA6 which was purified as described previously (see chapter 4). Previous data had confirmed the processing of parasite-derived SERA6 by rPfSUB1 (see chapter 4) and it was suggested that this proteolytic maturation frees SERA6 as an active protease. It was unknown which of the SERA6 forms might possess enzymatic activity. To not simply miss out on that activity by using the “wrong” form it was decided to use rPfSUB1-digested parasite-derived SERA6 samples that contain all SERA6 processing

fragments for activity screens. In preliminary experiments the processing of SERA6 was monitored by Western blot to identify the time required to digest the protein to completion. Figure 44A shows that 30 min of processing was sufficient to obtain preparations containing all SERA6 forms. Consequently samples from that time point were then used in activity screens. Processed SERA6 and appropriate controls (rPfSUB1 only, parasite-derived SERA6 only, and a buffer only control) were then incubated with several fluorogenic peptide substrates and fluorescence measurements taken at various time-points. Samples were digested to completion after 72 hours by addition of Pronase to reveal maximum fluorescence levels. Figure 44B shows that only one substrate, the TACE substrate IV (Abz-LAQAVRSSR-Dpa), showed a clear increase in fluorescence emission. The assay was then repeated with two different concentrations of the TACE IV substrate peptide. Measurements were taken after 5 min and overnight incubation at 37°C. This time no clear difference could be observed and several repetitions of the experiment delivered inconclusive results (see Figure 44C). These data suggested that this approach was not suitable for detecting parasite-derived SERA6 protease activity.

5.2.4 Activity of recombinant and soluble SERA6 requires further investigation

To continue investigation into SERA6 activity it was decided to screen the rS6C1 fragment from the soluble fraction of *E. coli* for putative proteolytic activity against fluorogenic peptide substrates. This was possible as, as described earlier, a small proportion of the expressed rS6C1 was expressed in the soluble fraction of recombinant *E. coli* cells. It was predicted that the rS6C1 found in the soluble fraction of *E. coli* cell lysates was correctly folded and therefore active. Since several attempts to purify it from the *E. coli* lysate failed (see chapter 3), it was decided to use the whole *E. coli* cell lysate containing rS6C1 for the activity screens in this section. As a control for these assays, a catalytic site mutant, rS6C1 C644A, was designed. If, after the incubation with fluorogenic substrate peptides, emission levels increased in the rS6C1 and not the mutant rS6C1 C644A samples, this should reflect SERA6 activity.

Both rS6C1 and rS6C1 C644A were recombinantly expressed in *E. coli* cells and a BugBuster Master Mix soluble fraction (crude bacterial lysate) was separated from IB by centrifugation. Fluorogenic peptide assays were then conducted with the crude bacterial lysate without further purification. Several AMC fluorogenic substrate peptides (see Table 11) were incubated with the lysates at 37°C and measurements were taken at 0 hours and 4 hours. As controls, rS6C1 and rS6C1 C644A were incubated with buffer only, AMC peptide 7 was incubated with buffer only and AMC peptide 6 was processed with pronase. As seen in Figure 45 some peptides underwent hydrolysis as indicated by an increase in fluorescence. However, this activity was

equally found in both the rS6C1 and the rS6C1 C644A digests, indicating that the activity was likely due to contaminating bacterial proteases. The results suggested that this approach was unsuitable to detect SERA6 protease activity.

5.2.5 PbSERA3: An active cysteine protease that requires PfSUB1 processing for activation

As no protease activity associated with SERA6 could be detected with all the above described approaches, the characterisation of a SERA6 orthologue was attempted. This was facilitated by the fact that it was previously shown that PbSERA3 undergoes proteolytic processing by PfSUB1 in a manner similar to that of SERA6. Determining protease activity of a SERA6 orthologue appeared the best approach to gain insights into *in vitro* and *in vivo* substrates and identify the role of SERA6 in the *P. falciparum* life-cycle. The aim of this section was therefore to search for protease activity mediated by PbSERA3.

5.2.5.1 PbSERA3 shows autocatalytic protease activity

As described in chapter 4, rPbSERA3 undergoes processing by rPfSUB1 at three sites. However, the rPbSERA3 WT appeared to undergo a final additional processing step (conversion from 48 kDa to 42 kDa) which was not observed for the rPbSERA3 C639A and therefore unlikely to be caused by rPfSUB1. As rPbSERA3 WT was expressed in a soluble form and therefore likely correctly folded it was suspected that this step might have been caused by autocatalytic processing of rPbSERA3. To investigate whether the conversion from approximately 48 kDa to 42 kDa was indeed autocatalytic the recombinant protein was digested with rPfSUB1 with or without the addition of E64 (rPbSERA3 WT +/-E64). E64 is an irreversible, potent and highly selective cysteine protease inhibitor which acts by covalently binding to the active thiol group of cysteine proteases. PfSUB1, a serine protease, cannot be inhibited with E64. Therefore E64 was expected to prevent the final processing step if it was autocatalytic. Samples taken at intervals during the digestions were analysed by Western blot with α-PbS3C1. As controls the rPbSERA3 WT was incubated alone or in presence of rPfSUB1 and rPD.

Figure 46A shows that both sets of samples (-/+ E64) showed largely similar maturation patterns. However, whereas the recombinant PbSERA3 WT -E64 underwent the previously observed maturation from the full-length protein through several intermediates to a prominent band visible at approximately 42 kDa, the recombinant PbSERA3 WT +E64 failed to be completely converted to the 42 kDa fragment and both 48 kDa and 42 kDa forms were detected with the same signal intensity. These results were consistent with the idea that the final conversion of

rPbSERA3 might be due to a cysteine protease which was sensitive to E64. Figure 46A also shows that in the absence of rPfSUB1 or in the presence of an rPfSUB1 inhibitor no processing was observed at all. This indicated that the E64-sensitive protease responsible for the final conversion only became active after the initial proteolytic maturation of rPbSERA3 WT by rPfSUB1. This raised the possibility that the protease responsible for the final processing step might be rPbSERA3 WT after activation by rPfSUB1.

Although the data so far implied that the last processing step was performed by rPbSERA3 only a mutation of the catalytic Cys in rPbSERA3 to abort any cysteine protease activity could provide an adequate control. If the final processing step was mediated by PbSERA3 itself, a difference would be expected to be detected between rPbSERA3 WT and C639A after processing with rPfSUB1. To assess this, a mutant in which the catalytic Cys residue of PbSERA3 was replaced with an Ala residue (PbSERA3 C639A) was recombinantly expressed and purified in a manner identical to the approach used for rPbSERA3 WT. Both proteins were then digested with rPfSUB1 and digests analysed by Western blot with the α-PbS3C1 antibodies. Figure 46B shows that both proteins undergo an identical processing pattern except for the final conversion step (lanes 1, 2, 6-8). Here, whereas the rPbSERA3 WT was converted to the final form of 42 kDa, the C639A mutant accumulated at 48 kDa. This result shows that modification of the putative catalytic Cys in rPbSERA3 C639A prevented the final conversion, strongly suggesting that the final processing step was due to an autocatalytic activity of rPbSERA3 WT. Importantly the results also implied that the rPbSERA3 WT required processing by rPfSUB1 to become proteolytically active.

Of course it can be argued that processing PbSERA3 by PfSUB1 does not reconstitute physiological events as both are derived from two different *Plasmodium* species. However, PbSERA3 is predicted to be processed at the same sites by PfSUB1 and PbSUB1 which was confirmed by *in vitro* processing of peptide substrates (Catherine Suarez, unpublished). As recombinant PbSUB1 was unavailable during this study, PfSUB1 was used throughout this work. The PbSERA3 activity assay will purely be applied to identify putative SERA6 substrates and not to explore the role of PbSERA3 in the *P. berghei* life-cycle, therefore not causing any conflict about the authenticity.

5.2.5.2 Autocatalytic processing of rPbSERA3 is E64 sensitive

E64 inhibits most cysteine proteases at a concentration of 10 μM and parasite egress can also be prevented when parasites are cultured in the presence of 10 μM E64 (Boyle *et al.*, 2010b). The results from the previous section showed that 20 μM E64 only partially prevented the final processing step of rPbSERA3 WT it was therefore

decided to investigate whether increasing E64 concentrations might fully inhibit the final conversion step. To test this, recombinant PbSERA3 WT was digested with rPfSUB1 in the presence of increasing concentrations of E64 and the digests analysed by Western Blot as previously. As shown in Figure 46B the signal for the 42 kDa processing fragment decreased with increasing E64 concentrations and at 100 μM was very faint compared to the signal at 48 kDa (lanes 3-5). In comparison digests of rPbSERA3 C639A produced only the 48 kDa fragment, with no sign of production of the 42 kDa form (lane 6). The rPbSERA3 C639A-derived 48 kDa product comigrated on SDS PAGE with the 48 kDa fragment that was detected in rPbSERA3 WT +100 μM E64 samples, indicating that both bands might resemble the same form of rPbSERA3. This provided more evidence that the final conversion is dependent on a cysteine protease activity and additionally that this step is mediated by autocatalytic processing of rPbSERA3.

To investigate the effect of other protease inhibitors on the autocatalytic processing of rPbSERA3 WT digests were preformed with or without three different protease inhibitors. The serine and cysteine protease inhibitors leupeptin and antipain were tested first. Samples of digests were analysed by Western blot with the α-PbS3C1 antibodies. As shown in Figure 47A leupeptin and antipain had no effect on the autocatalytic processing since the maturation patterns were the same as the sample not containing any protease inhibitors. Next it was attempted to analyse whether the processing could be inhibited with recombinant PbICP (rPbICP), a cysteine protease inhibitor secreted by *P. berghei* parasites (Rennenberg *et al.*, 2010). Digests performed with or without 1 μM rPbICP (a kind gift from Volker Heussler,) were analysed by Western blot as previously. Figure 47B shows that the rPbICP did not inhibit the autocatalytic processing of rPbSERA3 WT. However, this result was not entirely conclusive since the PbSERA3 antibody cross-reacted with the rPbICP and many unspecific bands were detected. This was probably due to the fact that both, the rPbSERA3 used to raise the antibodies and the rPbICP used in this experiment, were produced in *E. coli*.

In summary this section showed that the autocatalytic processing step of rPbSERA3 WT was sensitive to 100 μM E64 but not to leupeptin or antipain and possibly not PbICP.

5.3 Discussion

5.3.1 Exploring SERA6 activity in the parasite

SERA5 has long believed to be indispensable in the parasite asexual life-cycle, since several attempts to knockout the gene *in vitro* have failed (Miller *et al.*,

2002b;McCoubrie *et al.*, 2007). Recent approaches have dissected the importance of SERA5 in the parasite asexual life-cyle further and have shown that though SERA5 is indispensable it is not a proteolytic enzyme as the active site Ser can be replaced with Ala in the parasite (Robert Stallmach, unpublished). In this chapter the same approach was applied to a study of SERA6 as previous data also suggested the protein to be required for parasite survival (Miller *et al.*, 2002b). In transfection experiment aimed at attempted replacement of the active site Cys644 with an Ala, 8 out of 11 wild-type controls showed integration but none out of seven Cys to Ala replacements revealed single homologous recombination events. These results strongly imply that SERA6 is an active cysteine protease in the parasite and that this activity is indispensable for parasite survival. It remains possible that the Cys644 is required for an essential fold conferring correct function of the protein, for example that Cys644 may not be required for protease activity but for correct folding of the protein. SERA6 encodes 29 Cys residues in total and SERA5 encodes for 28 Cys residues and these Cys residues are positionally conserved between SERA5 and SERA6. SERA5 however does not encode for a Cys residue at the active site but is still correctly folded, suggesting that it is likely not to be required for protein fold but for protease activity. Therefore it appears unlikely that the disruption of the SERA6 Cys644 has detrimental effects on the protein fold rather than an effect on a protease activity.

The substitution of an active site residue in order to investigate protease activity has been widely used (Clark *et al.*, 1978;Baker *et al.*, 1993;Pei *et al.*, 1996). Despite Ala being an adequate choice for an active site substitution as it is a small amino acid with only the sulfhydryl group missing in comparison to the Cys residue, it would be intriguing to analyse whether the parasite can tolerate a substitution of the catalytic Cys residue with a Ser residue. This would be an intriguing approach as the Ser residue might confer a serine protease activity on SERA6.

5.3.2 Exploring SERA6 activity *in vitro*

At the beginning of this work it was hypothesised that SERA6 may require processing by PfSUB1 to be matured to an active protease. It was therefore decided to express SERA6 recombinantly in *E. coli* and then use the recombinant material in activity screens, such as zymograms, autocatalytic processing assays and fluorogenic substrate peptide assays. Several attempts to express various rSERA6 fragments in *E. coli* based expression systems proved unsuccessful, the insolubility of recombinant *E. coli* derived SERA6 being the first obstacle that had to be overcome in this part of the work. Previous work from other groups has shown that cysteine proteases such as falcipain 2 were shown to possess proteolytic activity after the proteins were refolded from material which was derived from solubilised IB (Shenai *et al.*, 2000). As rS6C1

was predicted to carry the protease domain of SERA6 it was attempted to refold this recombinant protein following expression in IB. The results from these refolding studies were inconclusive. It was difficult to assess the correct refolding stage of rS6C1 since activity could not be employed as a refolding measurement tool. Others have shown that correct refolding of proteases such as falcipain 2 was revealed because substrates (haemoglobin in case of falcipain 2) were known and if the recombinant material showed protease activity it was correctly refolded. Although it appeared that some of rS6C1 was refolded, protein refolding as a tool to obtain active SERA6 was disregarded in subsequent approaches.

Several attempts to determine SERA6 protease activity by screening bacterial cell lysates containing rS6C1, parasite-derived SERA6 or the parasite-derived SERA6-central domain with various techniques such as fluorogenic substrate peptide assays, zymograms, autocatalytic processing assays and activity assays on iRBC samples failed to detect any protease activity. In general all activity screens attempted in this work were complicated by two major challenges. Firstly, it was unknown whether any of the above-mentioned SERA6 fragments possessed protease activity. Secondly, any activity of any of the SERA6 fragments could have been masked by not having the correct substrate available.

New tools for the study of SERA6 activity were developed at the end of this work and will be applied in future approaches. The rSERA6 WT and rSERA6 C644A were successfully produced as soluble proteins by expression COS-7 cells. These two recombinant proteins can be used to screen for activity. This is especially important as the C644A mutant constitutes a very suitable control. The use of fluorogenic peptide substrate libraries for screening for SERA6 activity could provide a wide range of putative substrates. Additionally the COS-7 cell-derived rSERA6 can be screened in assays with iRBCs or RBC ghosts similar to an approach in which new PfSUB1 substrates were identified (Silmon de Monerri, 2010;Silmon de Monerri *et al.*, 2011).

One of the major obstacles in activity screens with parasite-derived SERA6 was obtaining the protein in sufficient amounts for screens. Coomassie Blue staining of the protein preparations revealed that even after two different large-scale protein purification steps SERA6 was not the most abundant protein in those fractions and was not even visible on the gel. Therefore for future approaches it will be necessary to improve the purification of parasite-derived SERA6. The integration of an epitope-tag into the genomic *sera6* gene could provide such a tool for specific purification of the protein. As an example of this, Stallmach *et al.* were successful in integrating a TAP-tag into the genomic *sera5* gene (Robert Stallmach, unpublished). The TAP-tag contains a tobacco etch virus (TEV) protease cleavage site which was used to release ”mature” SERA5 from an affinity column after purification from parasite cultures (Rigaut

et al., 1999;Puig *et al.*, 2001). A similar approach could be employed for SERA6 as a tool to obtain sufficient amounts for protease activity screens, provided that the tag did not interfere with this activity.

5.3.3 Cysteine protease activity of a SERA6 orthologue: promising insights

PbSERA3 was recombinantly expressed with the expectation that an active SERA6 orthologue could successfully be employed to identify SERA6 substrates. This was a particularly intriguing possibility because the subcellular localisation and expression patterns of PbSERA3 are well understood and it has been shown to be indispensable for *P. berghei* parasite survival (Schmidt-Christensen *et al.*, 2008;Putrianti *et al.*, 2009). Processing assays with the recombinant protein produced very promising results as differences were immediately observed in the processing profile of protein processed with rPfSUB1 in the presence and absence of E64. The difference between these two samples was that final PbSERA3 maturation appeared to be almost completely ablated in the presence of E64. This suggested the involvement of a cysteine protease in rPbSERA3 processing, but one that became active only after initial processing by rPfSUB1. Analysis of the rPbSERA3 C639A mutant then revealed that the replacement of Cys with Ala showed an identical phenotype as rPbSERA3 inhibited with E64. This identified the final processing step as being due to an autocatalytic activity of rPbSERA3, showing for the first time that a Cys-type SERA possesses cysteine protease activity. Additional attempts to prevent the autocatalytic processing of rPbSERA3 WT with leupeptin and antipain did not prevent the final conversion. Results from inhibition assays with rPbICP were inconclusive due to cross-reactivity of the α-PbS3C1 antibody. The structure of PbICP was recently identified which indicated that the SERAs are unlikely targets for this cysteine protease inhibitor and therefore it would not prevent the final autocatalytic conversion of PbSERA3 WT (Hansen *et al.*, 2011).

In order to investigate into rPbSERA3 activity it was attempted to precisely identify the autocatalytic processing site, with the intention of using that information to generate a fluorogenic substrate peptide based on this site. Confirming the proteolytic activity by cleaving synthetic substrates based on autoprocessing sites has been used for several proteases including PfSUB1 (Sajid *et al.*, 2000). A fluorogenic substrate based on the autoprocessing site is likely to also be a good substrate *in trans*. This step required N-terminal sequencing of the 42 kDa fragment by Edman degradation. For this, a protein concentration of at least 10 pmol had to be obtained (corresponding to a readily visible band on a Coomassie Blue stained polyacrylamide gel). However, throughout this work rPbSERA3 amounts were never sufficient for this step despite thorough investigation to improve the efficiency of insect cell expression system. Future

work will include the improvement of purification methods of rPbSERA3 to achieve this aim. In the meantime many additional attempts were made to explore the activity of rPbSERA3 by employing a variety of protease inhibitors to attempt to block autocatalytic processing. Exploration of the activity of rPbSERA3 is at the very beginning, and in future work, after gaining enough recombinant protein, it can be thoroughly analysed with a variety of different methods including screening of fluorogenic substrate peptide libraries, zymograms and incubation with parasite material (*P. berghei* and *P. falciparum*). Future work with rPbSERA3 could reveal substrates that can be applied to screen for SERA6 activity *in vitro* and in the parasite; for example a substrate peptide based on the autoprocessing site of PbSERA3 might be also processed by an active SERA6.

During this work it was also discovered that rPbSERA3 was processed by rPfSUB1 at a third, unpredicted site. N-terminal sequencing revealed that this site is an unfavoured rPfSUB1 site (Koussis, 2009;Koussis *et al.*, 2009;Silmon de Monerri, 2010;Silmon de Monerri *et al.*, 2011).

5.3.4 SERA6 cysteine protease activity in the parasite

P. falciparum SERA6 has yet to be shown to possess cysteine protease activity. However, the data presented here shows that the catalytic Cys appears indispensable for parasite survival *in vitro* and a SERA6 orthologue possesses cysteine protease activity, strongly suggesting that SERA6 is an active protease which is required for parasite survival. Of particular interest is the observation that the protease activity of the SERA6 orthologue is E64 sensitive. E64 is an irreversible inhibitor of cysteine proteases and has been shown in a number of studies to prevent egress (Salmon *et al.*, 2001;Greenbaum *et al.*, 2002;Boyle *et al.*, 2010b). Recent data suggest that cysteine proteases are involved in the breakdown of the EPM and not the PVM as parasites cultured in the presence of E64 were shown to rupture the PVM but not the EPM (Glushakova *et al.*, 2008). Other, conflicting data suggest that E64 has an effect on the breakdown of the PVM (Wickham *et al.*, 2003). Albeit contradictory, all the studies indicate the importance of cysteine proteases in the breakdown of the membranes surrounding the parasites. In this work 100 μM E64 was found to prevent autocatalytic processing of rPbSERA3. However, a concentration of only 10 μM E64 is sufficient to prevent parasite egress. Homology modelling of the SERA5 papain-like domain with the protease domains of falcipain 2 and falcipain 3 (Chrislaine Withers-Martinez, unpublished) has suggested that the entrance to the S2 pocket in SERA5 is restricted by the aromatic ring of the Tyr735 residue. This Tyr residue is conserved in SERA6 (Tyr783) and PbSERA3 (Tyr780). E64 requires binding to the S subsites to then covalently bind to the active site Cys (Varughese *et al.*, 1989). As the aromatic

ring of the Tyr residue in the predicted S2 pocket restricts the access of E64, the inhibitor is unlikely to interact with the catalytic Cys of PbSERA3. Therefore E64 is a rather low potent inhibitor of PbSERA3. As a result it would be predicted that high concentrations of E64 might be required to if at all inhibit rPbSERA3. This is consistent with the results in this section as concentrations ≥100 μM E64 were required to prevent autocatalytic processing of PbSERA3. This suggests on the other hand that SERA6 may not be the cysteine protease that is inhibited by E64 to prevent egress in parasite cultures, where concentrations of as little as 10 μM can block egress. However, in agreement with recent data that E64 inhibits breakdown of the EPM but not PVM it could be speculated that SERA6 is involved in a cascade leading to PVM rupture. Further exploration of this hypothesis requires the identification of SERA6 activity and its substrates in the parasite.

Egress of *P. falciparum* parasites is poorly understood and it is difficult to even speculate about putative SERA6 substrates. In recent years new insights into egress have been gained with the identification of proteases involved in this step. PfSUB1 is the best understood proteolytic player in egress but it performs a dual function as it also primes the merozoite surface for invasion (Yeoh *et al.*, 2007;Koussis *et al.*, 2009;Silmon de Monerri *et al.*, 2011). In this work SERA6 was shown to be processed by PfSUB1, releasing a central fragment carrying the putative protease domain. After being processed by PfSUB1 SERA6, perhaps now an active protease, may move on to play its part in a cascade leading to egress. Additionally it is has to mentioned that the lack of homologous recombination events in *in vitro P. falciparum* cultures not always confirms an essential biological role of the respective gene product in the human host or mosquito vector. In order to address this future work should include transfection complementation approaches. In these experiments dual transfections will be conducted with the homologues recombination integration construct carrying the for the parasite putative deleterious gene modifications (for example deletion of *sera6st2* or *Cys644Ala*). At the same time an episomal construct carrying the wild type *sera6* to “rescue” or complement the introduced mutation would be transfected. If the results observed during this study are due to the mutation of endogenous *sera6* it can be expected that the complementation approach would ensure parasite survival whereas parasites with no complementary plasmid would not survive.

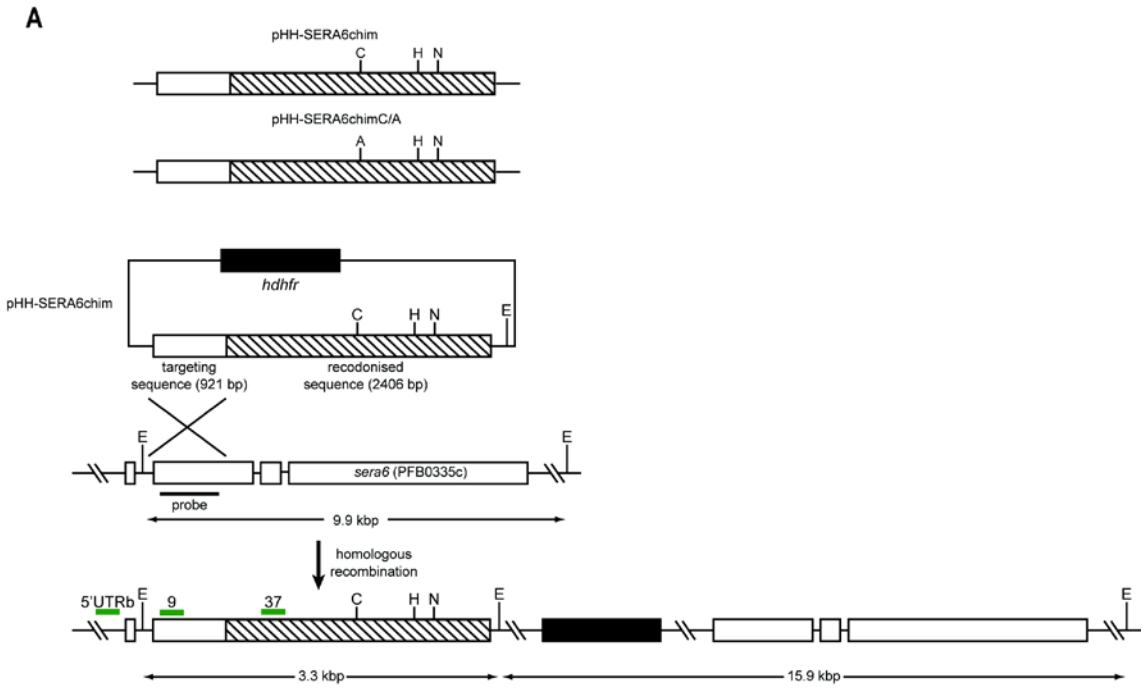
Future work could also include the modification of PkSERA4 in the *P. knowlesi* parasite. This might involve disrupting the entire gene, replacing the predicted SUB1 cleavage sites and replacing the active site Cys. If these modifications have similar effects to the modifications described for *sera6* this could provide another set of evidence for the importance of SERA6 and its orthologues.

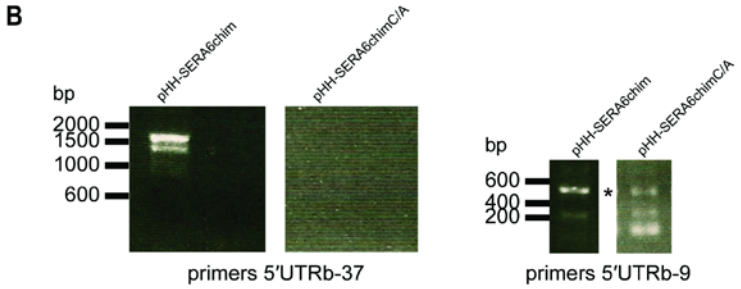
Table 15: Summary of all attempted SERA6 activity screens

SERA6 (origin)	***Method of activity screen***	***Activity***
E. coli refolded rS6-FL, rS6C1, rS6C2	Fluorogenic peptide substrate assays Autocatalytic processing Zymograms	Fluorogenic peptide assays positive due to contaminating proteases
E. coli soluble fraction rS6-FL, rS6C1, rS6C2	Fluorogenic peptide substrate assays Autocatalytic processing Zymograms	No
E. coli soluble fraction rS6C1 WT	Fluorogenic peptide substrate assays Autocatalytic processing	No
Parasite-derived semi-pure full-length SERA6	Fluorogenic peptide substrate assays Autocatalytic processing	No
Parasite-derived SERA6-central	Fluorogenic peptide substrate assays Autocatalytic processing Processing iRBC ghosts Zymograms	No
COS-7 cell derived rSERA6 WT	Autocatalytic processing	No
PbSERA3 (origin)	***Method of activity screen***	***Activity***
Tn5 insect cells rPbSERA3 WT	Autocatalytic processing	Yes

Figure 40: The predicted catalytic Cys of SERA6 is indispensable for parasite survival

Schematic of transfection strategy for single homologous recombination applied in this work (A). The strategy was based on a chimera between genomic *sera6* and recodonised *sera6* which facilitated single-homologous recombination in the targeted *sera6* locus. 3D7 parasites were transfected with two different constructs, pHH-SERA6chim and pHH-SERA6chimC/A to either reconstitute the wild-type SERA6 coding sequence or to replace the active site Cys644 with an Ala residue. Introduced mutations were located within the recodonised region of the transfection constructs. Transfected parasite lines were selected for integration by drug selection and genomic DNA was extracted after four drug cycles. Diagnostic PCR was applied to screen for integration. Positions of the primer pairs applied in the screens are indicated in green (A). The presence of genomic DNA was confirmed by using primer pair 5′ UTRb and 9 (B, right hand panel, asterisk). In a second PCR, with primer pair 5′ UTRb and 37, a band of the expected size (~1.8 kbp) was detected for pHH-SERA6chim but not for pHH-SERA6chimC/A indicating integration for the wild-type control but not the C644A mutant (B, left hand panel). A doublet was visible in all diagnostic PCRs of all transfected lines possibly due to the binding properties of the primer pair used. The uncloned PCR positive parasite lines were then analysed by Southern blot using the DIG method (C). A band at ~3.3 kbp was visible for parasites transfected with pHH-SERA6chim but not those transfected with pHH-SERA6chimC/A (C, red box). DNA of the transfection plasmid itself served as control in the PCR reactions (pHH-SERA6chimplasmid). Because the transfected lines were not clonal, both the input plasmid and the wild-type *sera6* locus were still detected by Southern blot (C, both ~9.5 kbp, asterisk). These results suggest that the *sera6* locus is accessible to homologous recombination but that substitution of the active site Cys residue may not be tolerated by the parasite.


A
pHH-SERA6chim
C
H N
pHH-SERA6chimC/A
A
H N
hdhfr
pHH-SERA6chim
C
H N
E
targeting
sequence (921 bp)
recodonised
sequence (2406 bp)
E
E
sera6 (PFB0335c)
probe
9.9 kbp
homologous
recombination
5'UTRb
E
9
37
C
H N
E
E
3.3 kbp
15.9 kbp


B
bp
pHH-SERA6chim
pHH-SERA6chimC/A
2000
1500
1000
600
primers 5'UTRb-37
bp
pHH-SERA6chim
pHH-SERA6chimC/A
600
400
200
*
primers 5'UTRb-9

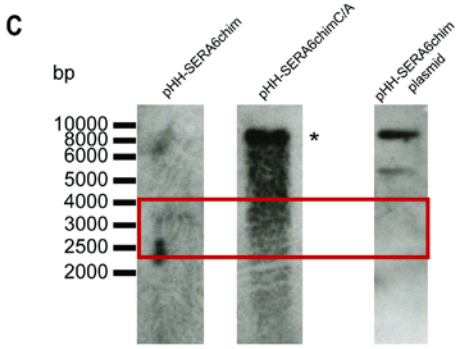

C
bp
pHH-SERA6chim
pHH-SERA6chimC/A
pHH-SERA6chim
plasmid
10000
8000
6000
5000
4000
3000
2500
2000
*

Figure 41: Refolded recombinant SERA6 showed no protease activity

Three different SERA6 fragments, rS6-FL, rS6C1 and rS6C2 were recombinantly expressed in *E. coli* IB. The IB were washed in 2 M urea buffer containing TX-100 before they were finally solubilised in 8 M urea. The preparations were then rapidly diluted 20 fold into a variety of refolding buffers and incubated for 4 days at 4°C. All samples were then analysed by CN-PAGE followed by Western blot. Western blots were probed with antibodies against the N-terminal 6xHis-tag. Samples refolded in buffer 1, 2, 3 (Quick Refolding Kit, Pierce), SERA5 refolding buffer and falcipain 2 refolding buffer were the only samples that showed bands on the CN-polyacrylamide gels, suggesting that these recombinant proteins were at least partially refolded (A). To next assess the level of refolding rS6-FL, rS6C1 and rS6C2 refolded in the falcipain 2 buffer, were concentrated, separated on SDS-polyacrylamide gels with or without DTT and stained with Coomassie Blue. The addition of DTT was expected to cause a size shift if the protein had formed intramolecular disulphide bonds during refolding. No clear size shift was observed (B). However, differences in protein amounts that entered the gel were visible for samples with or without DTT, suggesting refolding of a small proportion of recombinant protein (B). Samples of rS6C1 and rS6C2 refolded in falcipain 2 buffer were purified by size-exclusion chromatography and subjected to zymograms on gelatin gels, using papain activity buffer as developing buffer. No bands of clearing were detected, indicating an absence of activity in this assay. Activity was only observed for the papain control (C). IB: *E. coli* inclusion bodies; CN-PAGE: Clear Native PAGE

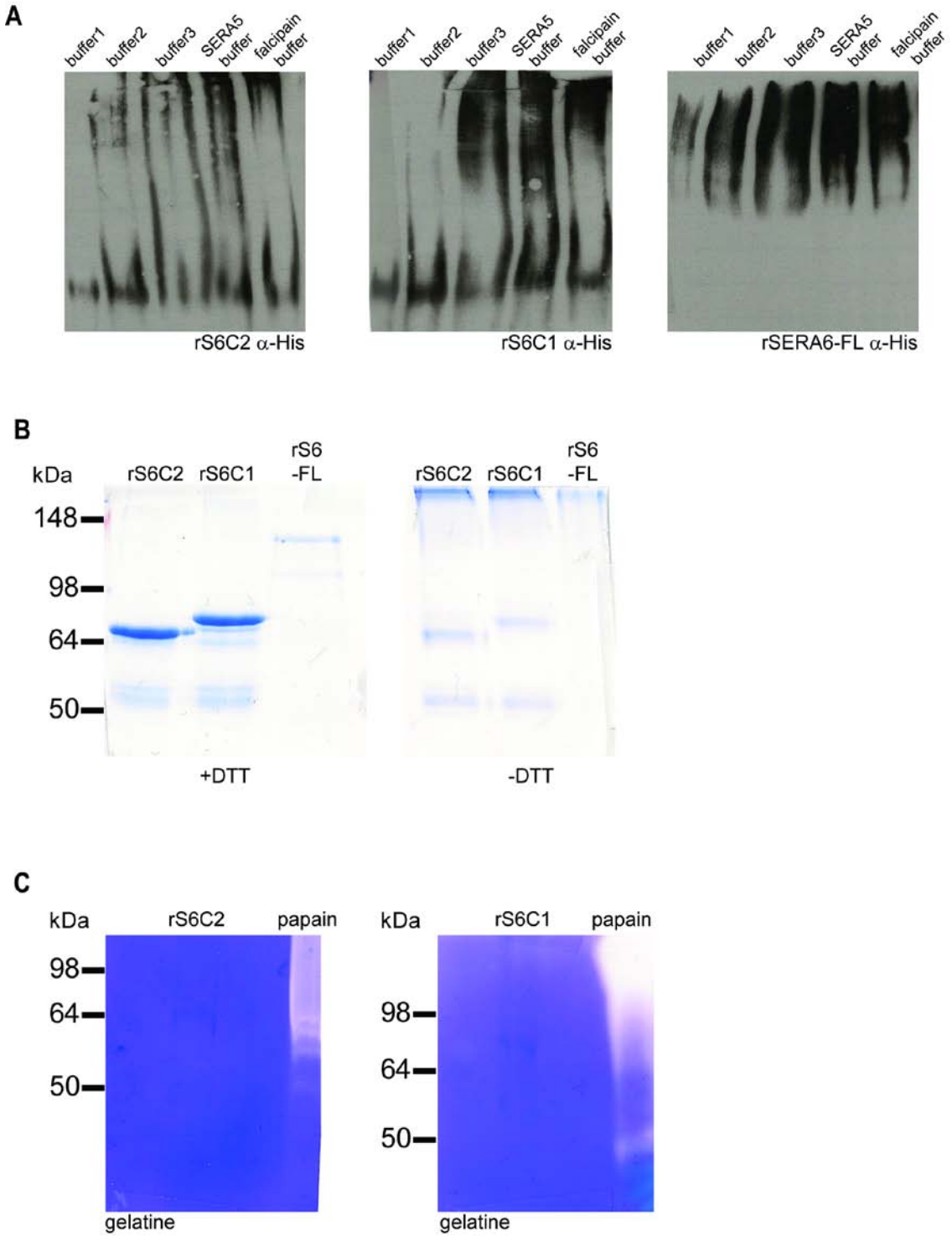
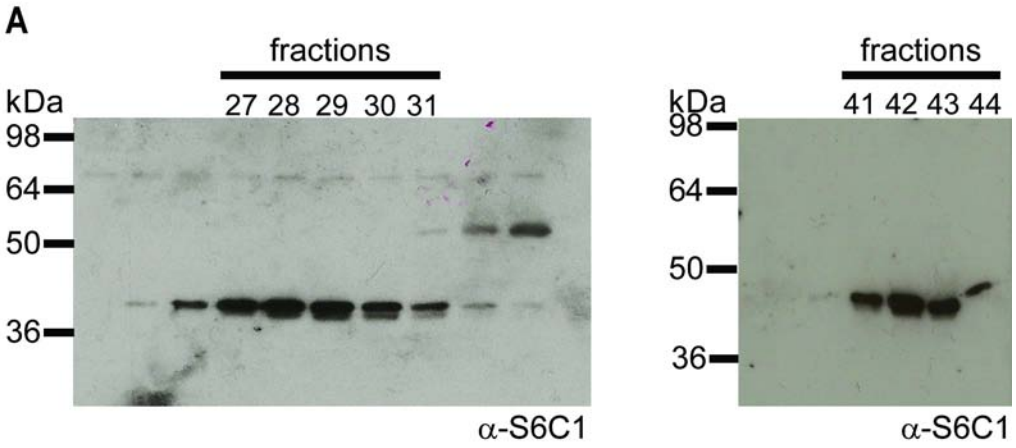

A
buffer1
buffer2
buffer3
SERA5 buffer
falcipain buffer
rS6C2 α-His
rS6C1 α-His
rSERA6-FL α-His
B
kDa
rS6C2
rS6C1
rS6 -FL
148
98
64
50
+DTT
-DTT
C
kDa
rS6C2
papain
rS6C1
98
64
50
gelatine

Figure 42: A protein corresponding to the SERA6 central domain is found in parasite culture supernatants after egress

Supernatants were collected after the egress of highly synchronous schizonts and fractionated by anion-exchange chromatography followed by size-exclusion chromatography. After each fractionation step collected samples were analysed by Western blot using the α-S6C1 antibody. After anion exchange chromatography a small fragment that reacted with the central SERA6 antibody was found in fractions 26 to 31 (A, left hand panel). These fractions were pooled and further purified by size-exclusion chromatography with a HiLoad 26/60 Superdex 200 prep-grade column. The antibody-reactive protein was now detected in fractions 41 to 44 (A, right hand panel). The peak fractions 42 and 43 were pooled, concentrated and stored frozen in aliquots. Samples were incubated at 37°C for 4 hours or overnight with and without Ca^{2+} and DTT, to assess for autocatalytic processing, and then analysed by SDS-PAGE and stained with Coomassie Blue staining (B). No differences between the start samples and the incubated samples were observed. The start samples were diluted 5 fold with SDS-sample buffer whereas the other samples were diluted 2 fold therefore the start sample appeared weaker after staining with Coomassie Blue. h: hours; o/n: overnight


A
fractions
kDa
27 28 29 30 31
98
64
50
36
α-S6C1
fractions
kDa
41 42 43 44
98
64
50
36
α-S6C1

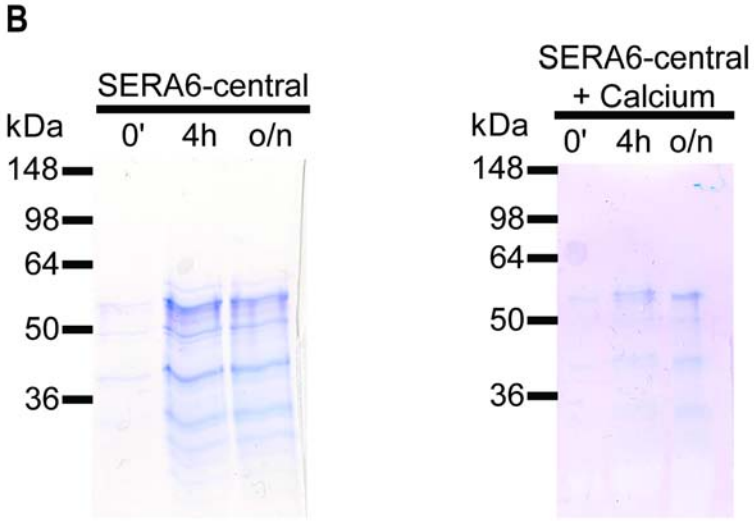

B
SERA6-central
kDa
0' 4h o/n
148
98
64
50
36
SERA6-central
+ Calcium
kDa
0' 4h o/n
148
98
64
50
36

Figure 43: Screen of SERA6-central detects no protease activity

The purified SERA6-central domain was screened for protease activity using two approaches. First it was incubated with iRBC ghosts with or without the addition of Ca^{2+}. The mix was incubated for 5 hours and samples were taken every hour, separated by SDS-PAGE and fractionated proteins visualised by silver stain. In order to check for any processing, samples were compared to the start sample (A). No differences were detected between the start sample and samples after 5 hours. Next SERA6-central was separated on polyacrylamide gels containing BSA as substrate to test for protease activity by zymograms. The zymograms were developed in four different buffers: falcipain 2 activity buffer, papain activity buffer, pH 8.2 activity buffer and pH 8.2 activity buffer with 5 mM $CaCl_2$. Papain was applied as positive control. As no bands of clearing were observed it was concluded that the SERA6-central domain was inactive or at least not processing the BSA substrate in any of the 4 different developing conditions (B). Papain was applied as a positive control. iRBC: infected red blood cells; h: hours

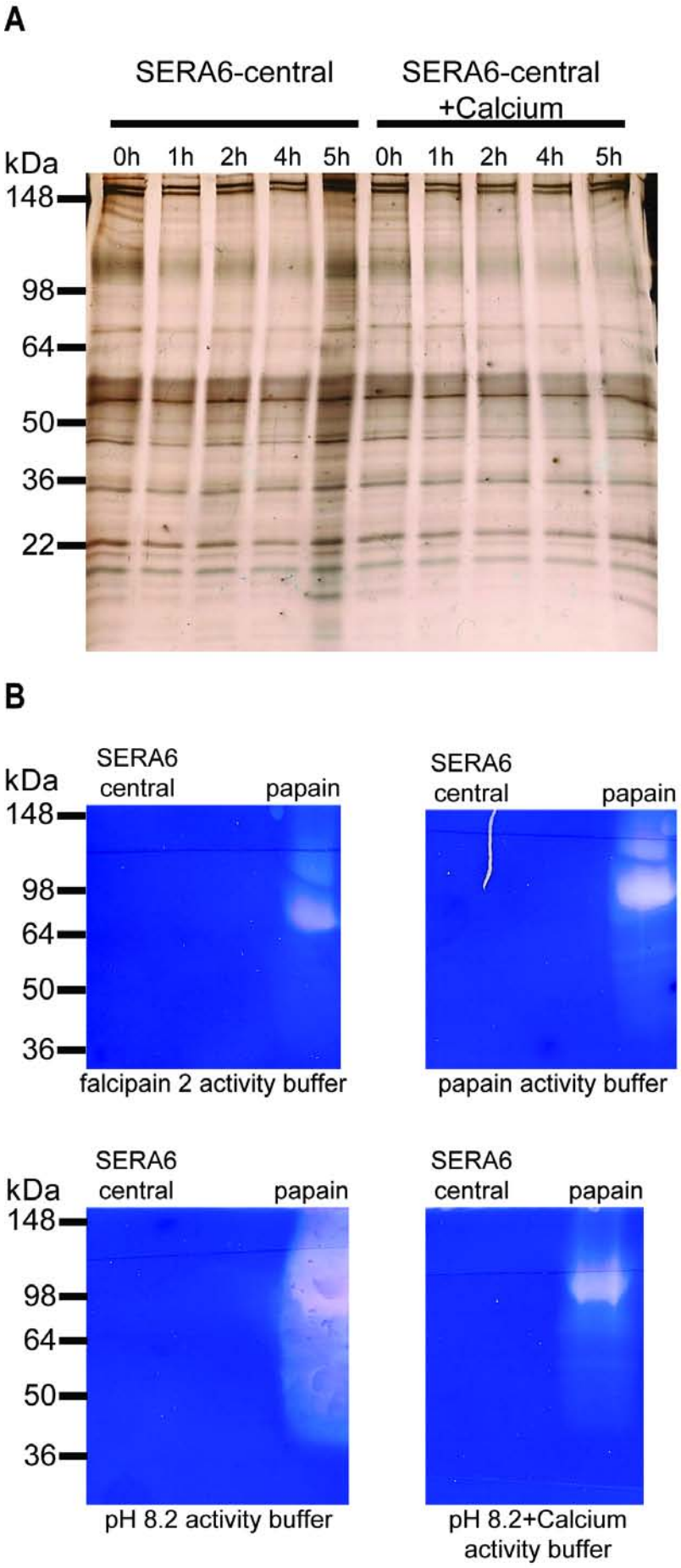

A
SERA6-central
SERA6-central +Calcium
kDa
0h 1h 2h 4h 5h 0h 1h 2h 4h 5h
148
98
64
50
36
22
B
kDa
SERA6 central
papain
148
98
64
50
36
falcipain 2 activity buffer
papain activity buffer
pH 8.2 activity buffer
pH 8.2+Calcium activity buffer

Figure 44: Parasite-derived SERA6 shows no activity in fluorogenic peptide substrate assays

Late stage schizonts were hypotonically lysed and full-length parasite SERA6 was purified by size-exclusion and anion-exchange chromatography. It was unknown which of the SERA6 fragments possessed protease activity therefore samples containing all SERA6 fragments were used in fluorogenic peptide substrate assays. Parasite-derived SERA6 was processed with rPfSUB1 and samples were analysed by Western blot using rabbit α-S6C1. SERA6 processing fragments appeared after 10 and 30 min (A). It was decided to use the 30 min sample for all subsequent activity screens (A, black box). A variety of fluorogenic peptide substrates was screened, with three controls included in the assays: buffer only, unprocessed SERA6 only and rPfSUB1 only (to test for any background processing). Samples were incubated for 5 min, 3 hours or 72 hours. After the final measurement pronase was added to cleave substrates to completion, resulting in maximum emission levels of each peptide. The TACE IV substrate (Abz-LAQAVRSSSR-Dpa) was the only peptide to show any increase in fluorescence after the incubation with processed parasite-derived SERA6 and was therefore the only one displayed below (B). The same peptide was then tested in triplicate at a concentration of 10 μM and 20 μM with several controls (SERA6 only, rPfSUB1 only and SERA6/rPfSUB1/rPD). This time the samples were incubated for 5 min or overnight and additional controls included SERA6 only, rPfSUB1 only and SERA6 processed with rPfSUB1 in the presence of the rPD. No difference in fluorescence levels was observed between the controls and the processed SERA6 samples. Hence it was suggested that the activity observed during the first assay was due to contaminating proteases. rPD: rPfSUB1 prodomain; o/n: overnight

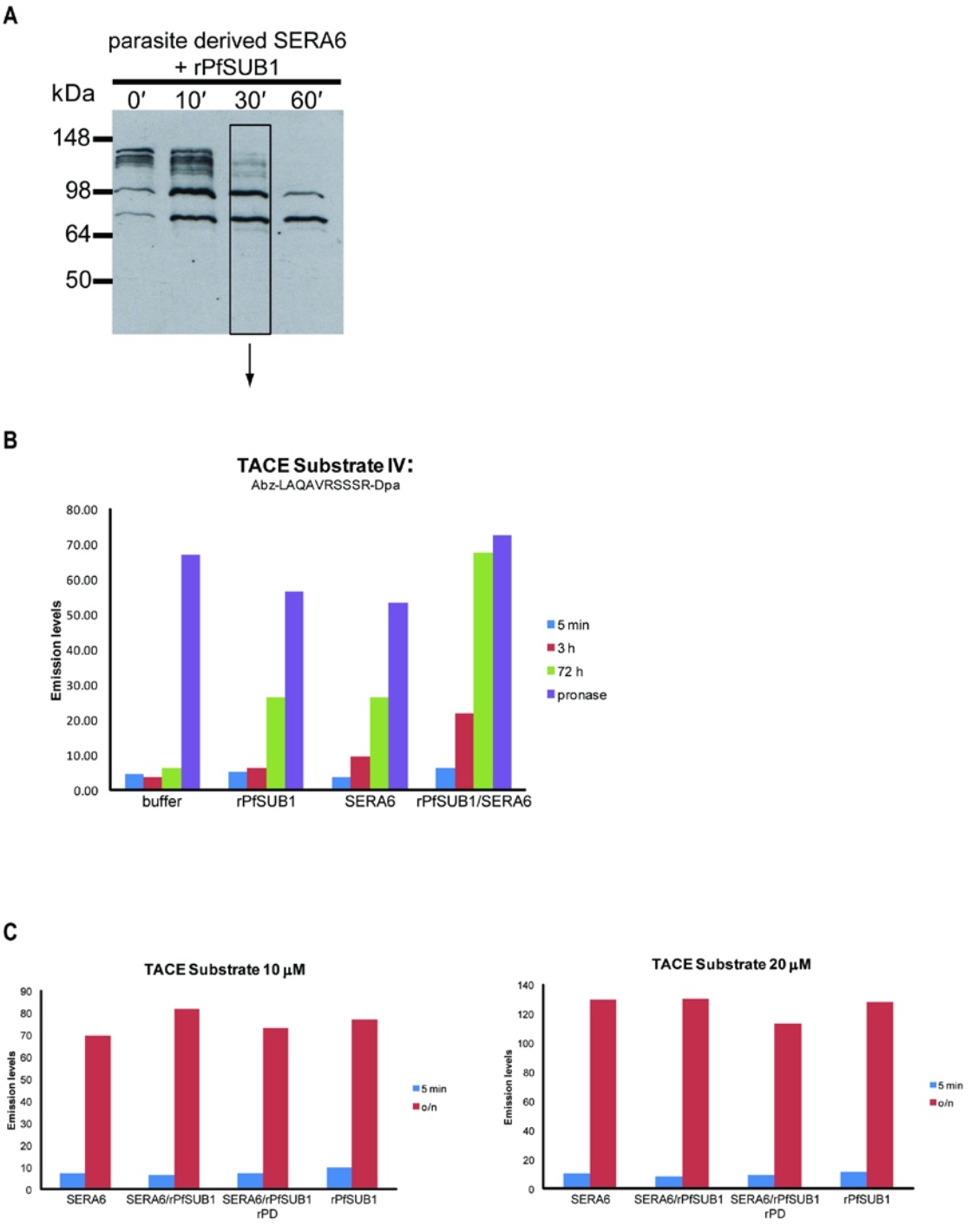

A
parasite derived SERA6
+ rPfSUB1
kDa
0′
10′
30′
60′
148
98
64
50
B
TACE Substrate IV:
Abz-LAQAVRSSSR-Dpa
Emission levels
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
buffer
rPfSUB1
SERA6
rPfSUB1/SERA6
5 min
3 h
72 h
pronase
C
TACE Substrate 10 µM
Emission levels
90
80
70
60
50
40
30
20
10
0
SERA6
SERA6/rPfSUB1
SERA6/rPfSUB1 rPD
rPfSUB1
5 min
o/n
TACE Substrate 20 µM
Emission levels
140
120
100
80
60
40
20
0
SERA6
SERA6/rPfSUB1
SERA6/rPfSUB1 rPD
rPfSUB1
5 min
o/n

Figure 45: *E. coli* lysates containing rS6C1 show no protease activity

To screen for protease activity *E. coli* expressing rS6C1 or rS6C1 C644A were collected and lysed with BugBuster Master Mix. The cell lysate was separated from IB by centrifugation. 15 AMC fluorogenic substrate peptides (numbers according to peptides displayed in Table 11 were resuspended in HEPES buffer and incubated with the *E. coli* cell lysate containing rS6C1 or rS6C1 C644A. The samples were incubated at 37°C and measurements were taken after 0 and 4 hours. Controls for the assay: rS6C1 and rS6C1 C644A with buffer only and 50 μg pronase with AMC peptide 6 after 4 hours. In order to detect activity of soluble rS6C1 differences in emission levels for each peptide after processing with either rS6C1 or rS6C1 C644A were required. None of the AMC peptides showed a clear difference between the wild-type and mutant rS6C1. h: hours; IB: *E. coli* inclusion bodies

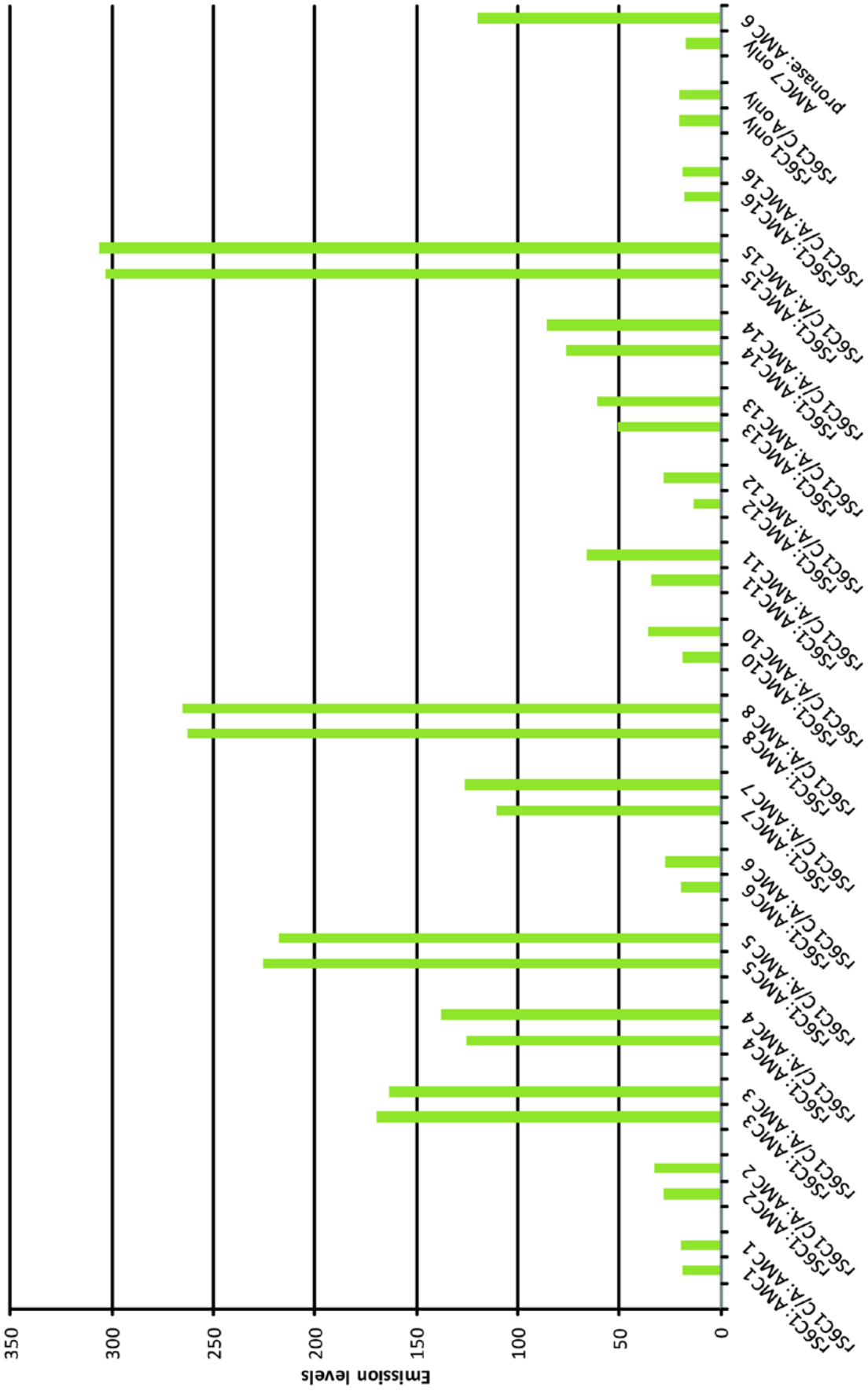
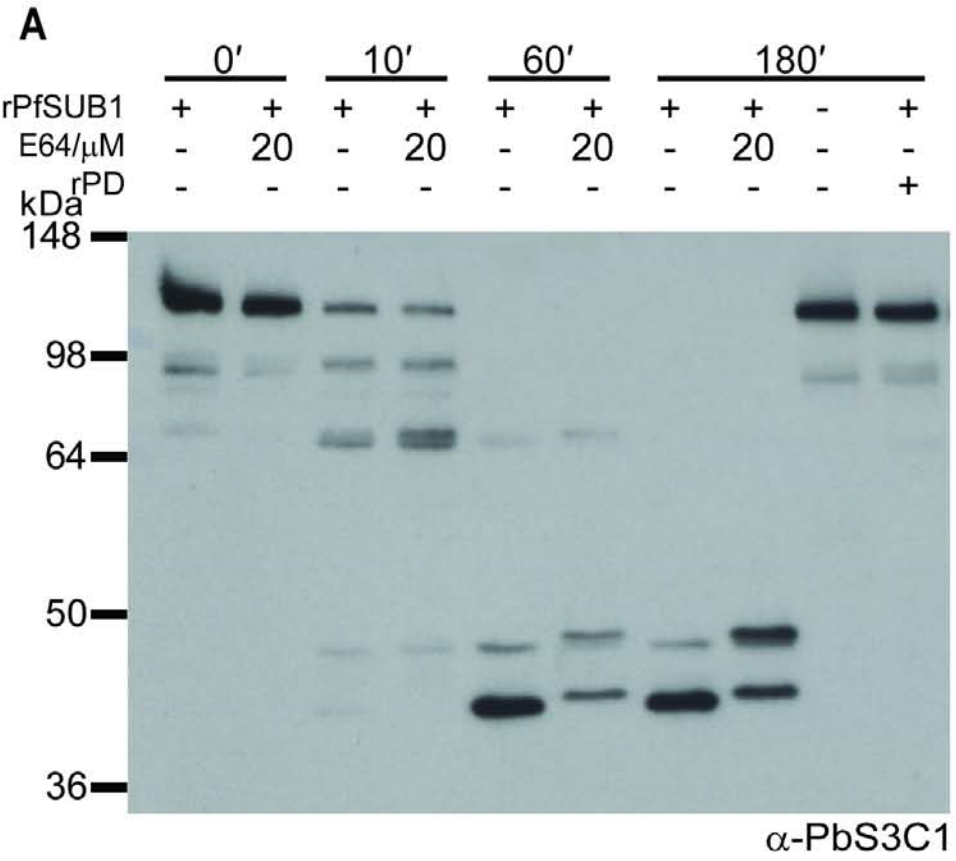

Emission levels
0
50
100
150
200
250
300
350
rSect C/A: AMC 1
rSect: AMC 1
rSect C/A: AMC 2
rSect: AMC 2
rSect C/A: AMC 3
rSect: AMC 3
rSect C/A: AMC 4
rSect: AMC 4
rSect C/A: AMC 5
rSect: AMC 5
rSect C/A: AMC 6
rSect: AMC 6
rSect C/A: AMC 7
rSect: AMC 7
rSect C/A: AMC 8
rSect: AMC 8
rSect C/A: AMC 10
rSect: AMC 10
rSect C/A: AMC 11
rSect: AMC 11
rSect C/A: AMC 12
rSect: AMC 12
rSect C/A: AMC 13
rSect: AMC 13
rSect C/A: AMC 14
rSect: AMC 14
rSect C/A: AMC 15
rSect: AMC 15
rSect C/A: AMC 16
rSect: AMC 16
rSect C/A only
rSect only
AMC 7 only
protease: AMC 6

Figure 46: Autocatalytic processing reveals PbSERA3 cysteine protease activity
Recombinant PbSERA3 WT was expressed and purified from insect cell supernatants and digested with rPfSUB1 with or without 20 μM E64. Controls included rPbSERA3 only and rPbSERA3 with rPfSUB1 and rPD. Samples were taken after 10, 60, 180 min and analysed by Western blot with α-PbS3C1 antibodies. In samples without E64 rPbSERA3 underwent almost complete processing and the majority accumulated at 42 kDa. In samples containing E64 the final conversion to the 42 kDa fragment appeared much slower and was never as prominent as in the rPbSERA3 without E64 sample. This suggests a role for a cysteine protease in the final conversion step. Both controls confirmed that the initial processing is mediated by rPfSUB1 as rPbSERA3 WT alone was stable as was rPbSERA3 with rPfSUB1 and rPD (A). Next it was investigated whether increased E64 concentrations could completely block the conversion to a 42 kDa fragment and more importantly whether the last processing step was due to autocatalytic processing of rPbSERA3 WT. To do this rPbSERA3 C639A was expressed and purified from insect cells precisely as described for the rPbSERA3 WT. Recombinant PbSERA3 WT was digested with rPfSUB1 in the presence of 10, 20 and 100 μM E64 and compared to rPbSERA3 C639A which was also processed with rPfSUB1. The conversion from 48 kDa to 42 kDa was almost completely inhibited by 100 μM E64 (B). The rPbSERA3 C639A mutant accumulated at the same 48 kDa form as the rPbSERA3 C639A. Both results implied that the final conversion to 42 kDa is due to autocatalytic processing of rPbSERA3 WT. rPD: rPfSUB1 prodomain; numbers below B refer to lanes


A
0′
10′
60′
180′
rPfSUB1 + + + + + + + + - +
E64/μM - 20 - 20 - 20 - 20 - -
rPD - - - - - - - - - +
kDa
148
98
64
50
36
α-PbS3C1

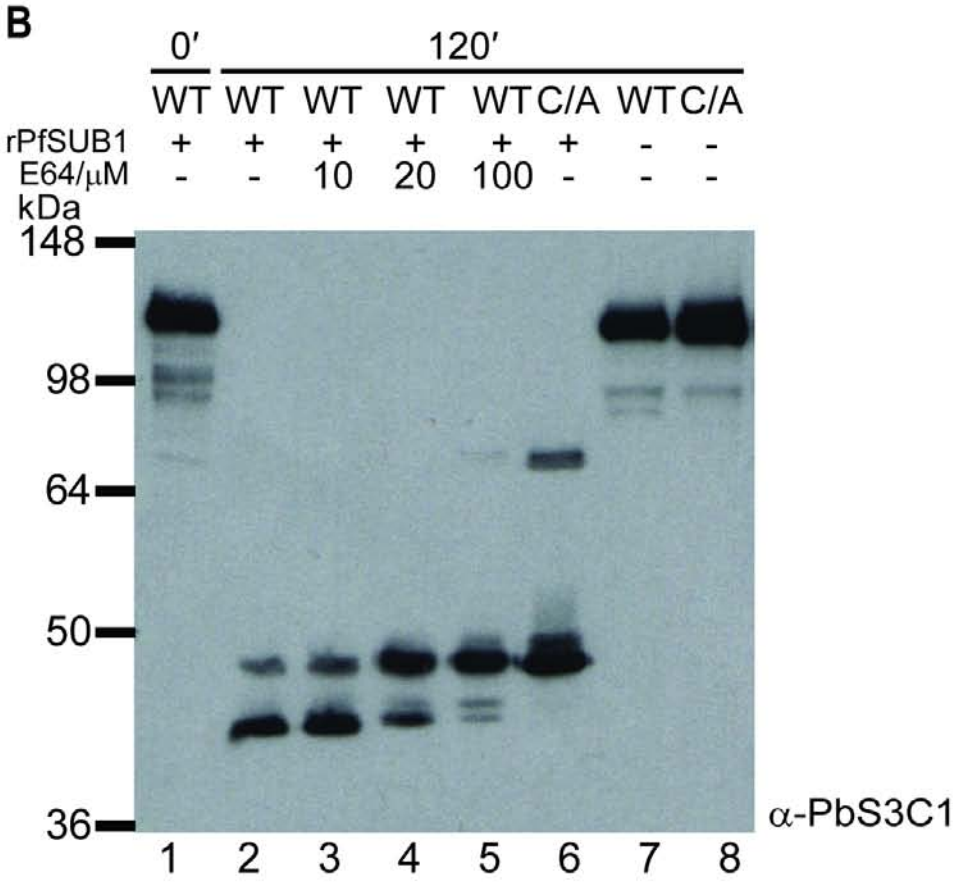
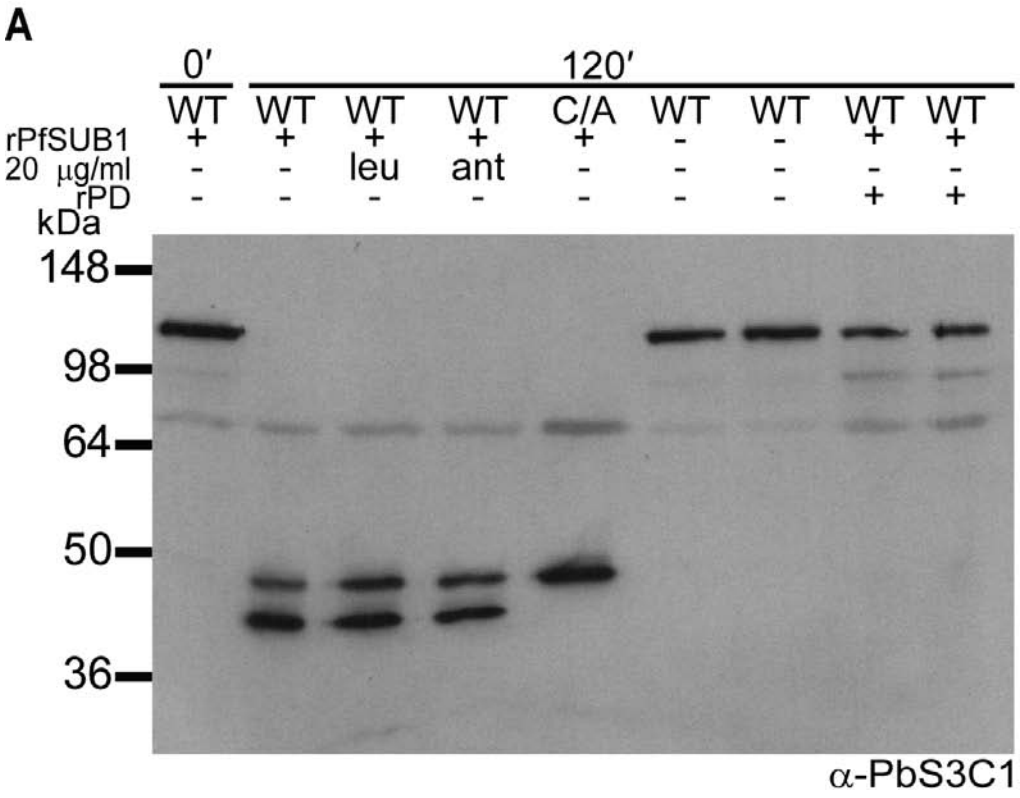
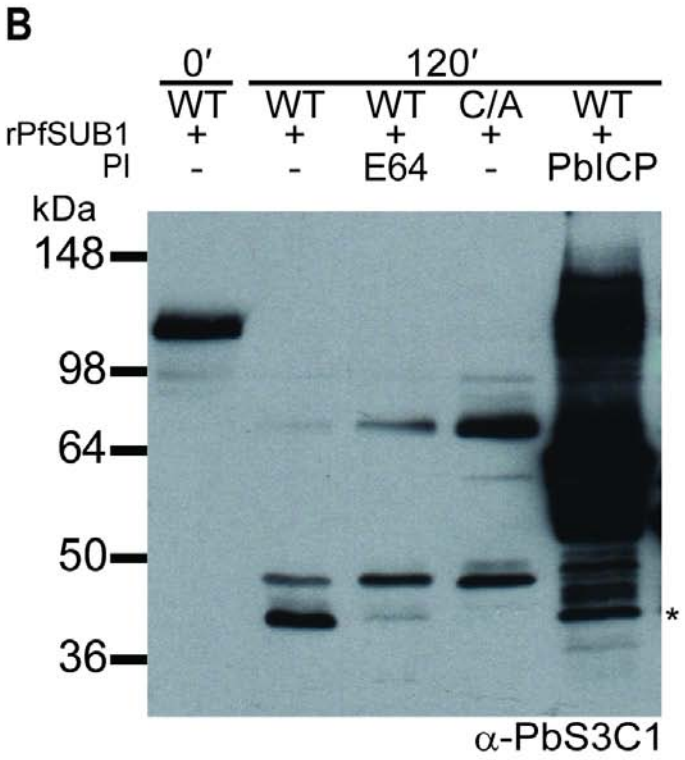

B
0′
120′
WT WT WT WT WT C/A WT C/A
rPfSUB1 + + + + + + - -
E64/μM - - 10 20 100 - - -
kDa
148
98
64
50
36
α-PbS3C1
1 2 3 4 5 6 7 8

Figure 47: Autocatalytic processing of PbSERA3 is insensitive to leupeptin, antipain, or PbICP

To test whether the autocatalytic processing can not only be prevented in the presence of E64, rPbSERA3 WT was digested with rPfSUB1 in the presence of the reversible cysteine and serine protease inhibitors leupeptin and antipain at a concentration of 20 μg/ml. After the incubation at 37°C samples were taken after 0 and 120 min and analysed by Western blot with α-PbS3C1 antibodies. Controls included rPbSERA3 WT and rPbSERA3 C639A only and in presence of both rPfSUB1 and rPD. Leupeptin and antipain had no effect on the autocatalytic processing of rPbSERA3 as samples containing the inhibitors showed identical processing patterns as samples without protease inhibitors (A). Next it was tested whether rPbICP, a potent cysteine protease inhibitor expressed by *P. berghei*, blocked the autocatalytic processing of rPbSERA3 WT. Controls included rPbSERA3 WT processed with rPfSUB1 in presence of 100 μM E64 and rPbSERA3 C639A processed with rPfSUB1. Samples were taken after 0 and 120 min and analysed by Western blot. Despite strong cross-reactivity between rPbICP and the α-PbS3C1 antibody the accumulation of the bottom 42 kDa band was still detectable (B, asterisk). Due to cross-reactivity of the antibodies it remained inconclusive whether the 42 kDa band in fact represented the rPbSERA3 final fragment showing that PbICP had no effect on the autocatalytic processing. rPD: rPfSUB1 prodomain; PI: protease inhibitor; leu: leupeptin; ant: antipain


A
0′
120′
WT WT WT WT C/A WT WT WT WT
rPfSUB1 + + + + + - - + +
20 µg/ml - - leu ant - - - - -
rPD - - - - - - - + +
kDa
148
98
64
50
36
α-PbS3C1


B
0′
120′
WT WT WT C/A WT
rPfSUB1 + + + + +
PI - - E64 - PbICP
kDa
148
98
64
50
36
*
α-PbS3C1

6 Conclusions and future perspectives: SERA6 and its role in the malaria life-cycle

At the beginning of this work SERA6 was predicted to be localised to the PV of late stage schizonts and some data was available showing that SERA6 encodes for processing sites that are cleavable by PfSUB1 *in vitro*. Most importantly previous work from other groups had suggested that SERA6 was important for blood stage parasites. Therefore this study attempted to explore (A) the localisation of SERA6 (B) the processing of SERA6 by PfSUB1 and finally (C) the protease activity of SERA6. Five main conclusions can now be drawn from the work in this study.

1. SERA6 is located to the PV of *in vitro P. falciparum* blood stage schizonts
2. SERA6 is processed by PfSUB1 late in the asexual life-cycle *in vitro*
3. Part of the PfSUB1 processing of SERA6 appears to be important for the parasite *in vitro*
4. The catalytic Cys of SERA6 may be required for parasite survival *in vitro*
5. PbSERA3, a SERA6 orthologue, becomes an active cysteine protease upon processing by PfSUB1

It can therefore be hypothesised that SERA6 is an essential cysteine protease that becomes active upon proteolytic maturation by PfSUB1 late in the intra-erythrocytic stages of *P. falciparum in vitro* cultures. The question remains however, what is the exact role of SERA6 in the parasite life-cycle. Based on results from this study this last section will explore the possibility of SERA6 being involved in a proteolytic cascade leading to egress.

Many cysteine proteases are known to become active enzymes upon proteolytic maturation (Voet *et al.*, 2004). Once activated through processing by PfSUB1, SERA6 might cleave its substrates in the PV or the host cell cytosol. During localisation experiments in this work it was noted that a small proportion of SERA6 is membrane associated. It is unknown whether this membrane is derived from the PVM or EPM. Does this membrane association imply that upon activation SERA6 cleaves integral or peripheral associated membrane proteins triggering the breakdown of the PVM and the EPM? PV or PVM proteins interacting with SERA6 could include EXP1 and the early transcribed membrane proteins (ETRAMPs), which are integral proteins of the PVM. EXP1 and ETRAMPs are thought to form pores in the PVM which might be important in maintaining the structural integrity of this membrane (Spielmann *et al.*, 2006). Upon proteolytic processing perhaps by SERA6, EXP1 and some or all members of the ETRAMP family could undergo structural changes causing the breakdown of the PVM.

This breakdown might also be triggered by another interaction pathway. SERA6 might also process a member of the *Plasmodium* perforin-like protein family. Perforin-like proteins are soluble proteins that can form a membrane pore complex upon physiological changes such as proteolytic maturation (Uellner *et al.*, 1997;Kafsack *et al.*, 2010). Evidence for a role for the perforin-like proteins in the breakdown of membranes surrounding the parasite and a role in egress has been given recently (Ecker *et al.*, 2007;Kafsack *et al.*, 2008;Kafsack *et al.*, 2010). Therefore processing of a perforin-like protein by SERA6 might cause structural changes to the protein that enable it to play its role in egress by inducing pores in the PVM or EPM.

The involvement of SERA6 in the breakdown of the PVM is also supported by observations made in IFAs. These results showed a strong SERA6 signal in late stage schizonts with tightly packed merozoites. These tightly packed merozoites arguably indicate a still intact PVM. In IFAs of schizonts with widely distributed merozoites but intact EPM no SERA6 signal was observed. These schizonts possibly displayed a ruptured PVM, suggesting that the rupture of the PVM coincides with the loss of SERA6. These observations possibly suggest a SERA6 role in the breakdown of the PVM.

SERA6 might have several roles throughout egress, which may not only include the breakdown of the PVM but also the EPM. The RBC cytoskeleton proteins α-spectrin, ankyrin, band 4.9 and actin are thought to undergo proteolytic processing around egress (Millholland *et al.*, 2011). Therefore upon rupture of the PVM, active SERA6 may translocate into the RBC cytosol where it may process proteins of the RBC cytoskeleton. This is supported by the experimental evidence that E64 prevents the rupture of the EPM but not the PVM (Wickham *et al.*, 2003). Hence the breakdown of the EPM appears cysteine protease dependent and the protease involved might be SERA6 if SERA6 can be inhibited with E64. Other proteases, such as falcipain 2, plasmepsin II or the host cell calpain, may be involved in rupture of the EPM (Le Bonniec *et al.*, 1999;Dua *et al.*, 2001;Hanspal *et al.*, 2002;Chandramohanadas *et al.*, 2009;Silmon de Monerri, 2010;Millholland *et al.*, 2011).

SERA5 has previously been considered to possess weak proteolytic activity but recent data suggest that SERA5 is essential but not an essential enzyme (Robert Stallmach, unpublished). Due to the results shown in this work, it can be proposed that SERA5 and SERA6 have distinct roles in the parasite. In order to analyse the protease activity of SERA6 future work will require the screen for PbSERA3 substrates which will eventually lead to the identification of putative SERA6 substrates. Any putative substrates could then be used to identify the activity of SERA6 *in vitro* and in the parasite.

In summary results from this work showed that SERA6 is located to the PV where it undergoes proteolytic maturation by PfSUB1, possibly releasing SERA6 as an active cysteine protease which is essential for parasite survival. Proteases have been implicated in many diseases and serve as excellent drug targets and play key roles in the treatment of for example HIV/AIDS and cancer (Podgorski *et al.*, 2003;Drag *et al.*, 2010). SERA6 is unique to *Plasmodium* and no homologues can be found in the human host which makes it an excellent target for the development of new antimalarials (Venter *et al.*, 2001;Arisue *et al.*, 2011).

Bibliography

(1979). Antimalaria studies on Qinghaosu. *Chin Med J (Engl)* **92,** 811-816.

Aaslyng, D., Gormsen, E. and Malmos, H. (1991). Mechanistic Studies of Proteases and Lipases for the Detergent Industry. *J Chem Technol Biot* **50,** 321-330.

Abkarian, M., Massiera, G., Berry, L., Roques, M. and Braun-Breton, C. (2011). A novel mechanism for egress of malarial parasites from red blood cells. *Blood* **117,** 4118-4124.

Aikawa, M. (1988). Human cerebral malaria. *Am J Trop Med Hyg* **39,** 3-10.

Aikawa, M. and Miller, L.H. (1983). Structural alteration of the erythrocyte membrane during malarial parasite invasion and intraerythrocytic development. *Ciba Found Symp* **94,** 45-63.

Aikawa, M., Miller, L.H., Johnson, J. and Rabbege, J. (1978). Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol* **77,** 72-82.

Aikawa, M., Torii, M., Sjolander, A., Berzins, K., Perlmann, P. and Miller, L.H. (1990). Pf155/RESA antigen is localized in dense granules of Plasmodium falciparum merozoites. *Exp Parasitol* **71,** 326-329.

Alexander, D.L., Mital, J., Ward, G.E., Bradley, P. and Boothroyd, J.C. (2005). Identification of the moving junction complex of Toxoplasma gondii: a collaboration between distinct secretory organelles. *PLoS Pathog* **1,** e17.

Allison, A.C. (1954). Protection afforded by sickle-cell trait against subtertian malareal infection. *Br Med J* **1,** 290-294.

Alonso, P.L., Sacarlal, J., Aponte, J.J., Leach, A., Macete, E., Aide, P.*, et al.* (2005). Duration of protection with RTS,S/AS02A malaria vaccine in prevention of Plasmodium falciparum disease in Mozambican children: single-blind extended follow-up of a randomised controlled trial. *Lancet* **366,** 2012-2018.

Aly, A.S. and Matuschewski, K. (2005). A malarial cysteine protease is necessary for Plasmodium sporozoite egress from oocysts. *J Exp Med* **202,** 225-230.

Anderson, E.D., Molloy, S.S., Jean, F., Fei, H., Shimamura, S. and Thomas, G. (2002). The ordered and compartment-specfific autoproteolytic removal of the furin intramolecular chaperone is required for enzyme activation. *J Biol Chem* **277,** 12879-12890.

Anderson, E.D., VanSlyke, J.K., Thulin, C.D., Jean, F. and Thomas, G. (1997). Activation of the furin endoprotease is a multiple-step process: requirements for acidification and internal propeptide cleavage. *EMBO J* **16,** 1508-1518.

Ansorge, I., Benting, J., Bhakdi, S. and Lingelbach, K. (1996). Protein sorting in Plasmodium falciparum-infected red blood cells permeabilized with the pore-forming protein streptolysin O. *Biochem J* **315 (Pt 1),** 307-314.

Ansorge, I., Paprotka, K., Bhakdi, S. and Lingelbach, K. (1997). Permeabilization of the erythrocyte membrane with streptolysin O allows access to the vacuolar membrane of Plasmodium falciparum and a molecular analysis of membrane topology. *Mol Biochem Parasitol* **84,** 259-261.

Aoki, S., Li, J., Itagaki, S., Okech, B.A., Egwang, T.G., Matsuoka, H.*, et al.* (2002). Serine repeat antigen (SERA5) is predominantly expressed among the SERA multigene family of Plasmodium falciparum, and the acquired antibody titers correlate with serum inhibition of the parasite growth. *J Biol Chem* **277,** 47533-47540.

Arastu-Kapur, S., Ponder, E.L., Fonovic, U.P., Yeoh, S., Yuan, F., Fonovic, M.*, et al.* (2008). Identification of proteases that regulate erythrocyte rupture by the malaria parasite Plasmodium falciparum. *Nat Chem Biol* **4,** 203-213.

Arisue, N., Hirai, M., Arai, M., Matsuoka, H. and Horii, T. (2007). Phylogeny and evolution of the SERA multigene family in the genus Plasmodium. *J Mol Evol* **65,** 82-91.

Arisue, N., Kawai, S., Hirai, M., Palacpac, N.M., Jia, M., Kaneko, A.*, et al.* (2011). Clues to evolution of the SERA multigene family in 18 Plasmodium species. *PLoS ONE* **6,** e17775.

Armstrong, C.M. and Goldberg, D.E. (2007). An FKBP destabilization domain modulates protein levels in Plasmodium falciparum. *Nat Methods*.

Bailly, E., Jambou, R., Savel, J. and Jaureguiberry, G. (1992). Plasmodium falciparum: differential sensitivity in vitro to E-64 (cysteine protease inhibitor) and Pepstatin A (aspartyl protease inhibitor). *J Protozool* **39,** 593-599.

Baker, D. and Sali, A. (2001). Protein structure prediction and structural genomics. *Science* **294,** 93-96.

Baker, R.P., Wijetilaka, R. and Urban, S. (2006). Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog* **2,** e113.

Baker, S.C., Yokomori, K., Dong, S., Carlisle, R., Gorbalenya, A.E., Koonin, E.V. and Lai, M.M. (1993). Identification of the catalytic sites of a papain-like cysteine proteinase of murine coronavirus. *J Virol* **67,** 6056-6063.

Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M. and Goldberg, D.E. (2002). Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. *Proc Natl Acad Sci U S A* **99,** 990-995.

Bannister, L. and Mitchell, G. (2003). The ins, outs and roundabouts of malaria. *Trends Parasitol* **19,** 209-213.

Bannister, L.H., Butcher, G.A., Dennis, E.D. and Mitchell, G.H. (1975). Structure and invasive behaviour of Plasmodium knowlesi merozoites in vitro. *Parasitology* **71,** 483-491.

Bannister, L.H. and Dluzewski, A.R. (1990). The ultrastructure of red cell invasion in malaria infections: a review. *Blood Cells* **16,** 257-292; discussion 293-257.

Barale, J.C., Attal-Bonnefoy, G., Brahimi, K., Pereira da Silva, L. and Langsley, G. (1997). Plasmodium falciparum asparagine and aspartate rich protein 2 is an evolutionary conserved protein whose repeats identify a new family of parasite antigens. *Mol Biochem Parasitol* **87,** 169-181.

Barale, J.C., Blisnick, T., Fujioka, H., Alzari, P.M., Aikawa, M., Braun-Breton, C. and Langsley, G. (1999). Plasmodium falciparum subtilisin-like protease 2, a merozoite candidate for the merozoite surface protein 1-42 maturase. *Proc Natl Acad Sci U S A* **96,** 6445-6450.

Barrett, A.J. and Rawlings, N.D. (1995). Families and clans of serine peptidases. *Arch Biochem Biophys* **318,** 247-250.

Barrett, A.J. and Rawlings, N.D. (2001). Evolutionary lines of cysteine peptidases. *Biol Chem* **382,** 727-733.

Baum, J., Gilberger, T.W., Frischknecht, F. and Meissner, M. (2008). Host-cell invasion by malaria parasites: insights from Plasmodium and Toxoplasma. *Trends Parasitol* **24,** 557-563.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.W*., et al.* (2006). A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. *J Biol Chem* **281,** 5197-5208.

Baumeister, S., Winterberg, M., Duranton, C., Huber, S.M., Lang, F., Kirk, K. and Lingelbach, K. (2006). Evidence for the involvement of Plasmodium falciparum proteins in the formation of new permeability pathways in the erythrocyte membrane. *Mol Microbiol* **60,** 493-504.

Bender, M.L. and Killheffer, J.V. (1973). Chymotrypsins. *CRC Crit Rev Biochem* **1,** 149-199.

Berry, C., Humphreys, M.J., Matharu, P., Granger, R., Horrocks, P., Moon, R.P*., et al.* (1999). A distinct member of the aspartic proteinase gene family from the human malaria parasite Plasmodium falciparum. *FEBS Lett* **447,** 149-154.

Besteiro, S., Dubremetz, J.F. and Lebrun, M. (2011). The moving junction of apicomplexan parasites: a key structure for invasion. *Cell Microbiol*.

Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.F. and Lebrun, M. (2009). Export of a Toxoplasma gondii rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion. *PLoS Pathog* **5,** e1000309.

Billker, O., Dechamps, S., Tewari, R., Wenig, G., Franke-Fayard, B. and Brinkmann, V. (2004). Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* **117,** 503-514.

Blackman, M.J. (2000). Proteases involved in erythrocyte invasion by the malaria parasite: function and potential as chemotherapeutic targets. *Curr Drug Targets* **1,** 59-83.

Blackman, M.J. (2004). Proteases in host cell invasion by the malaria parasite. *Cell Microbiol* **6,** 893-903.

Blackman, M.J. (2008). Malarial proteases and host cell egress: an 'emerging' cascade. *Cell Microbiol* **10,** 1925-1934.

Blackman, M.J., Chappel, J.A., Shai, S. and Holder, A.A. (1993). A conserved parasite serine protease processes the Plasmodium falciparum merozoite surface protein-1. *Mol Biochem Parasitol* **62,** 103-114.

Blackman, M.J., Corrie, J.E., Croney, J.C., Kelly, G., Eccleston, J.F. and Jameson, D.M. (2002). Structural and biochemical characterization of a fluorogenic rhodamine-labeled malarial protease substrate. *Biochemistry* **41,** 12244-12252.

Blackman, M.J., Fujioka, H., Stafford, W.H., Sajid, M., Clough, B., Fleck, S.L*., et al.* (1998). A subtilisin-like protein in secretory organelles of Plasmodium falciparum merozoites. *J Biol Chem* **273,** 23398-23409.

Boddey, J.A., Hodder, A.N., Gunther, S., Gilson, P.R., Patsiouras, H., Kapp, E.A*., et al.* (2009). An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463,** 627-631.

Bonilla, J.A., Bonilla, T.D., Yowell, C.A., Fujioka, H. and Dame, J.B. (2007a). Critical roles for the digestive vacuole plasmepsins of Plasmodium falciparum in vacuolar function. *Mol Microbiol* **65,** 64-75.

Bonilla, J.A., Moura, P.A., Bonilla, T.D., Yowell, C.A., Fidock, D.A. and Dame, J.B. (2007b). Effects on growth, hemoglobin metabolism and paralogous gene expression resulting from disruption of genes encoding the digestive vacuole plasmepsins of Plasmodium falciparum. *Int J Parasitol* **37,** 317-327.

Bourgon, R., Delorenzi, M., Sargeant, T., Hodder, A.N., Crabb, B.S. and Speed, T.P. (2004). The serine repeat antigen (SERA) gene family phylogeny in Plasmodium: the impact of GC content and reconciliation of gene and species trees. *Mol Biol Evol* **21,** 2161-2171.

Boyle, M.J., Richards, J.S., Gilson, P.R., Chai, W. and Beeson, J.G. (2010a). Interactions with heparin-like molecules during erythrocyte invasion by Plasmodium falciparum merozoites. *Blood* **115,** 4559-4568.

Boyle, M.J., Wilson, D.W., Richards, J.S., Riglar, D.T., Tetteh, K.K., Conway, D.J*., et al.* (2010b). Isolation of viable Plasmodium falciparum merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proc Natl Acad Sci U S A* **107,** 14378-14383.

Brier, B. (2004). Infectious diseases in ancient Egypt. *Infect Dis Clin North Am* **18,** 17-27.

Brooker, S., Guyatt, H., Omumbo, J., Shretta, R., Drake, L. and Ouma, J. (2000). Situation analysis of malaria in school-aged children in Kenya - what can be done? *Parasitol Today* **16,** 183-186.

Brossier, F., Jewett, T.J., Lovett, J.L. and Sibley, L.D. (2003). C-terminal processing of the toxoplasma protein MIC2 is essential for invasion into host cells. *J Biol Chem* **278,** 6229-6234.

Brossier, F., Jewett, T.J., Sibley, L.D. and Urban, S. (2005). A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by Toxoplasma. *Proc Natl Acad Sci U S A* **102,** 4146-4151.

Brossier, F., Starnes, G.L., Beatty, W.L. and Sibley, L.D. (2008). Microneme rhomboid protease TgROM1 is required for efficient intracellular growth of Toxoplasma gondii. *Eukaryot Cell* **7,** 664-674.

Buckle, A.M., Devlin, G.L., Jodun, R.A., Fulton, K.F., Faux, N., Whisstock, J.C. and Bottomley, S.P. (2005). The matrix refolded. *Nat Methods* **2,** 3.

Bull, P.C., Berriman, M., Kyes, S., Quail, M.A., Hall, N., Kortok, M.M*., et al.* (2005). Plasmodium falciparum variant surface antigen expression patterns during malaria. *PLoS Pathog* **1,** e26.

Bzik, D.J., Li, W.B., Horii, T. and Inselburg, J. (1988). Amino acid sequence of the serine-repeat antigen (SERA) of Plasmodium falciparum determined from cloned cDNA. *Mol Biochem Parasitol* **30,** 279-288.

Cabrita, L.D. and Bottomley, S.P. (2004). Protein expression and refolding--a practical guide to getting the most out of inclusion bodies. *Biotechnol Annu Rev* **10,** 31-50.

Cao, J., Kaneko, O., Thongkukiatkul, A., Tachibana, M., Otsuki, H., Gao, Q*., et al.* (2009a). Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in Plasmodium falciparum merozoites. *Parasitol Int* **58,** 29-35.

Cao, Y., Zhang, D. and Pan, W. (2009b). Construction of transgenic Plasmodium berghei as a model for evaluation of blood-stage vaccine candidate of Plasmodium falciparum chimeric protein 2.9. *PLoS ONE* **4,** e6894.

Cappadoro, M., Giribaldi, G., O'Brien, E., Turrini, F., Mannu, F., Ulliers, D*., et al.* (1998). Early phagocytosis of glucose-6-phosphate dehydrogenase (G6PD)-deficient erythrocytes parasitized by Plasmodium falciparum may explain malaria protection in G6PD deficiency. *Blood* **92,** 2527-2534.

Carruthers, V.B., Giddings, O.K. and Sibley, L.D. (1999a). Secretion of micronemal proteins is associated with toxoplasma invasion of host cells. *Cell Microbiol* **1,** 225-235.

Carruthers, V.B., Sherman, G.D. and Sibley, L.D. (2000). The Toxoplasma adhesive protein MIC2 is proteolytically processed at multiple sites by two parasite-derived proteases. *J Biol Chem* **275,** 14346-14353.

Carruthers, V.B. and Sibley, L.D. (1999b). Mobilization of intracellular calcium stimulates microneme discharge in Toxoplasma gondii. *Mol Microbiol* **31,** 421-428.

Carter, P. and Wells, J.A. (1988). Dissecting the catalytic triad of a serine protease. *Nature* **332,** 564-568.

Carter, R. and Mendis, K.N. (2002). Evolutionary and historical aspects of the burden of malaria. *Clin Microbiol Rev* **15,** 564-594.

Casares, S., Brumeanu, T.D. and Richie, T.L. (2010). The RTS,S malaria vaccine. *Vaccine* **28,** 4880-4894.

Catteruccia, F. (2007). Malaria vector control in the third millennium: progress and perspectives of molecular approaches. *Pest Manag Sci* **63,** 634-640.

Chakravorty, S.J., Hughes, K.R. and Craig, A.G. (2008). Host response to cytoadherence in Plasmodium falciparum. *Biochem Soc Trans* **36,** 221-228.

Chandramohanadas, R., Davis, P.H., Beiting, D.P., Harbut, M.B., Darling, C., Velmourougane, G.*, et al.* (2009). Apicomplexan Parasites Co-Opt Host Calpains to Facilitate Their Escape from Infected Cells. *Science*.

Child, M.A., Epp, C., Bujard, H. and Blackman, M.J. (2010). Regulated maturation of malaria merozoite surface protein-1 is essential for parasite growth. *Mol Microbiol* **78,** 187-202.

Chow, M.K., Amin, A.A., Fulton, K.F., Whisstock, J.C., Buckle, A.M. and Bottomley, S.P. (2006). REFOLD: an analytical database of protein refolding methods. *Protein Expr Purif* **46,** 166-171.

Christophides, G.K. (2005). Transgenic mosquitoes and malaria transmission. *Cell Microbiol* **7,** 325-333.

Chulay, J.D., Lyon, J.A., Haynes, J.D., Meierovics, A.I., Atkinson, C.T. and Aikawa, M. (1987). Monoclonal antibody characterization of Plasmodium falciparum antigens in immune complexes formed when schizonts rupture in the presence of immune serum. *J Immunol* **139,** 2768-2774.

Ciechanover, A. (2011). Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Biochim Biophys Acta*.

Clague, M.J. and Urbe, S. (2010). Ubiquitin: same molecule, different degradation pathways. *Cell* **143,** 682-685.

Clark, I.A. and Cowden, W.B. (2003). The pathophysiology of falciparum malaria. *Pharmacol Ther* **99,** 221-260.

Clark, P.I. and Lowe, G. (1978). Conversion of the active-site cysteine residue of papain into a dehydro-serine, a serine and a glycine residue. *Eur J Biochem* **84,** 293-299.

Clavijo, C.A., Mora, C.A. and Winograd, E. (1998). Identification of novel membrane structures in Plasmodium falciparum infected erythrocytes. *Mem Inst Oswaldo Cruz* **93,** 115-120.

Coatney, G.R. (1963). Pitfalls in a discovery: the chronicle of chloroquine. *Am J Trop Med Hyg* **12,** 121-128.

Coldren, R.L., Jongsakul, K., Vayakornvichit, S., Noedl, H. and Fukudas, M.M. (2007). Apparent relapse of imported Plasmodium ovale malaria in a pregnant woman. *Am J Trop Med Hyg* **77,** 992-994.

Collins, C.R., Withers-Martinez, C., Hackett, F. and Blackman, M.J. (2009). An inhibitory antibody blocks interactions between components of the malarial invasion machinery. *PLoS Pathog* **5,** e1000273.

Collins, W.E. and Jeffery, G.M. (2005). Plasmodium ovale: parasite and disease. *Clin Microbiol Rev* **18,** 570-581.

Coloma, M.J., Hastings, A., Wims, L.A. and Morrison, S.L. (1992). Novel vectors for the expression of antibody molecules using variable regions generated by polymerase chain reaction. *J Immunol Methods* **152,** 89-104.

Conway, D.J. and Baum, J. (2002). In the blood--the remarkable ancestry of Plasmodium falciparum. *Trends Parasitol* **18,** 351-355.

Conway, D.J., Fanello, C., Lloyd, J.M., Al-Joubori, B.M., Baloch, A.H., Somanath, S.D.*, et al.* (2000). Origin of Plasmodium falciparum malaria is traced by mitochondrial DNA. *Mol Biochem Parasitol* **111,** 163-171.

Coombs, G.H., Goldberg, D.E., Klemba, M., Berry, C., Kay, J. and Mottram, J.C. (2001). Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. *Trends Parasitol* **17,** 532-537.

Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J.S. and Cygler, M. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J* **15,** 5492-5503.

Cowman, A.F. and Crabb, B.S. (2006). Invasion of red blood cells by malaria parasites. *Cell* **124,** 755-766.

Cox-Singh, J., Davis, T.M., Lee, K.S., Shamsul, S.S., Matusop, A., Ratnam, S.*, et al.* (2008a). Plasmodium knowlesi malaria in humans is widely distributed and potentially life threatening. *Clin Infect Dis* **46,** 165-171.

Cox-Singh, J. and Singh, B. (2008b). Knowlesi malaria: newly emergent and of public health importance? *Trends Parasitol* **24,** 406-410.

Cox, F.E. (2010). History of the discovery of the malaria parasites and their vectors. *Parasit Vectors* **3,** 5.

Crabb, B.S. (2002). Transfection technology and the study of drug resistance in the malaria parasite Plasmodium falciparum. *Drug Resist Updat* **5,** 126-130.

Craig, M.H., Snow, R.W. and le Sueur, D. (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol Today* **15,** 105-111.

Crompton, P.D., Pierce, S.K. and Miller, L.H. (2010). Advances and challenges in malaria vaccine development. *J Clin Invest* **120,** 4168-4178.

Daily, J.P. (2006). Antimalarial drug therapy: the role of parasite biology and drug resistance. *J Clin Pharmacol* **46,** 1487-1497.

Dawes, E.J., Churcher, T.S., Zhuang, S., Sinden, R.E. and Basanez, M.G. (2009). Anopheles mortality is both age- and Plasmodium-density dependent: implications for malaria transmission. *Malar J* **8,** 228.

Debrabant, A., Maes, P., Delplace, P., Dubremetz, J.F., Tartar, A. and Camus, D. (1992). Intramolecular mapping of Plasmodium falciparum P126 proteolytic fragments by N-terminal amino acid sequencing. *Mol Biochem Parasitol* **53,** 89-95.

Deguercy, A., Hommel, M. and Schrevel, J. (1990). Purification and characterization of 37-kilodalton proteases from Plasmodium falciparum and Plasmodium berghei which cleave erythrocyte cytoskeletal components. *Mol Biochem Parasitol* **38,** 233-244.

Delplace, P., Bhatia, A., Cagnard, M., Camus, D., Colombet, G., Debrabant, A.*, et al.* (1988). Protein p126: a parasitophorous vacuole antigen associated with the release of Plasmodium falciparum merozoites. *Biol Cell* **64,** 215-221.

Delplace, P., Dubremetz, J.F., Fortier, B. and Vernes, A. (1985). A 50 kilodalton exoantigen specific to the merozoite release-reinvasion stage of Plasmodium falciparum. *Mol Biochem Parasitol* **17,** 239-251.

Delplace, P., Fortier, B., Tronchin, G., Dubremetz, J.F. and Vernes, A. (1987). Localization, biosynthesis, processing and isolation of a major 126 kDa antigen of the parasitophorous vacuole of Plasmodium falciparum. *Mol Biochem Parasitol* **23,** 193-201.

Dhawan, S., Dua, M., Chishti, A.H. and Hanspal, M. (2003). Ankyrin peptide blocks falcipain-2-mediated malaria parasite release from red blood cells. *J Biol Chem* **278,** 30180-30186.

Dluzewski, A.R., Ling, I.T., Rangachari, K., Bates, P.A. and Wilson, R.J. (1984). A simple method for isolating viable mature parasites of Plasmodium falciparum from cultures. *Trans R Soc Trop Med Hyg* **78,** 622-624.

Dluzewski, A.R., Rangachari, K., Wilson, R.J. and Gratzer, W.B. (1986). Plasmodium falciparum: protease inhibitors and inhibition of erythrocyte invasion. *Exp Parasitol* **62,** 416-422.

Doby, J.M. and Guiguen, C. (1981). [2 Breton cases of autochthonous malaria, in reality the first French cases of "airport malaria"]. *Bull Soc Pathol Exot Filiales* **74,** 398-405.

Dowse, T.J., Koussis, K., Blackman, M.J. and Soldati-Favre, D. (2008). Roles of proteases during invasion and egress by Plasmodium and Toxoplasma. *Subcell Biochem* **47,** 121-139.

Dowse, T.J. and Soldati, D. (2005). Rhomboid-like proteins in Apicomplexa: phylogeny and nomenclature. *Trends Parasitol* **21,** 254-258.

Drag, M. and Salvesen, G.S. (2010). Emerging principles in protease-based drug discovery. *Nature reviews* **9,** 690-701.

Drew, M.E., Banerjee, R., Uffman, E.W., Gilbertson, S., Rosenthal, P.J. and Goldberg, D.E. (2008). Plasmodium food vacuole plasmepsins are activated by falcipains. *J Biol Chem* **283,** 12870-12876.

Dua, M., Raphael, P., Sijwali, P.S., Rosenthal, P.J. and Hanspal, M. (2001). Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. *Mol Biochem Parasitol* **116,** 95-99.

Duraisingh, M.T., Triglia, T., Ralph, S.A., Rayner, J.C., Barnwell, J.W., McFadden, G.I. and Cowman, A.F. (2003). Phenotypic variation of Plasmodium falciparum merozoite proteins directs receptor targeting for invasion of human erythrocytes. *EMBO J* **22,** 1047-1057.

Dvorak, J.A., Miller, L.H., Whitehouse, W.C. and Shiroishi, T. (1975). Invasion of erythrocytes by malaria merozoites. *Science* **187,** 748-750.

Dvorin, J.D., Martyn, D.C., Patel, S.D., Grimley, J.S., Collins, C.R., Hopp, C.S.*, et al.* (2010). A plant-like kinase in Plasmodium falciparum regulates parasite egress from erythrocytes. *Science* **328,** 910-912.

Ecker, A., Bushell, E.S., Tewari, R. and Sinden, R.E. (2008). Reverse genetics screen identifies six proteins important for malaria development in the mosquito. *Mol Microbiol* **70,** 209-220.

Ecker, A., Pinto, S.B., Baker, K.W., Kafatos, F.C. and Sinden, R.E. (2007). Plasmodium berghei: plasmodium perforin-like protein 5 is required for mosquito midgut invasion in Anopheles stephensi. *Exp Parasitol* **116,** 504-508.

Edman, P. (1949). A method for the determination of amino acid sequence in peptides. *Arch Biochem* **22,** 475.

Eggleson, K.K., Duffin, K.L. and Goldberg, D.E. (1999). Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum. *J Biol Chem* **274,** 32411-32417.

Eksi, S. and Williamson, K.C. (2011). Protein targeting to the Parasitophorous vacuole membrane of Plasmodium falciparum. *Eukaryot Cell* **10,** 744-752.

Ellis, R.D., Martin, L.B., Shaffer, D., Long, C.A., Miura, K., Fay, M.P.*, et al.* (2010). Phase 1 trial of the Plasmodium falciparum blood stage vaccine MSP1(42)-C1/Alhydrogel with and without CPG 7909 in malaria naive adults. *PLoS ONE* **5,** e8787.

Endo, T., Sethi, K.K. and Piekarski, G. (1982). Toxoplasma gondii: calcium ionophore A23187-mediated exit of trophozoites from infected murine macrophages. *Exp Parasitol* **53,** 179-188.

Escalante, A.A. and Ayala, F.J. (1994). Phylogeny of the malarial genus Plasmodium, derived from rRNA gene sequences. *Proc Natl Acad Sci U S A* **91,** 11373-11377.

Escalante, A.A., Barrio, E. and Ayala, F.J. (1995). Evolutionary origin of human and primate malarias: evidence from the circumsporozoite protein gene. *Mol Biol Evol* **12,** 616-626.

Fang, X.D., Kaslow, D.C., Adams, J.H. and Miller, L.H. (1991). Cloning of the Plasmodium vivax Duffy receptor. *Mol Biochem Parasitol* **44,** 125-132.

Fidock, D.A., Nomura, T., Talley, A.K., Cooper, R.A., Dzekunov, S.M., Ferdig, M.T.*, et al.* (2000). Mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell* **6,** 861-871.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D.*, et al.* (2002). A proteomic view of the Plasmodium falciparum life cycle. *Nature* **419,** 520-526.

Fox, B.A., Xing-Li, P., Suzue, K., Horii, T. and Bzik, D.J. (1997). Plasmodium falciparum: an epitope within a highly conserved region of the 47-kDa amino-terminal domain of the serine repeat antigen is a target of parasite-inhibitory antibodies. *Exp Parasitol* **85,** 121-134.

Francischetti, I.M., Seydel, K.B. and Monteiro, R.Q. (2008). Blood coagulation, inflammation, and malaria. *Microcirculation* **15,** 81-107.

Freeman, M. (2008). Rhomboid proteases and their biological functions. *Annu Rev Genet* **42,** 191-210.

Freeman, M. (2009). Rhomboids: 7 years of a new protease family. *Seminars in cell & developmental biology* **20,** 231-239.

Galinski, M.R., Medina, C.C., Ingravallo, P. and Barnwell, J.W. (1992). A reticulocyte-binding protein complex of Plasmodium vivax merozoites. *Cell* **69,** 1213-1226.

Gallup, J.L. and Sachs, J.D. (2001). The economic burden of malaria. *Am J Trop Med Hyg* **64,** 85-96.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W.*, et al.* (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419,** 498-511.

Garnham, P.C., Lainson, R. and Cooper, W. (1957). The tissue stages and sporogony of Plasmodium knowlesi. *Trans R Soc Trop Med Hyg* **51,** 384-396.

Garnham, P.C.C. (1966) Malaria parasites and other haemosporidia Oxford,, Blackwell Scientific, pp. xviii, 1114 p.

Gelhaus, C., Fritsch, J., Krause, E. and Leippe, M. (2005). Fractionation and identification of proteins by 2-DE and MS: towards a proteomic analysis of Plasmodium falciparum. *Proteomics* **5,** 4213-4222.

Ghosh, A.K. and Jacobs-Lorena, M. (2009). Plasmodium sporozoite invasion of the mosquito salivary gland. *Curr Opin Microbiol* **12,** 394-400.

Gilson, P.R. and Crabb, B.S. (2009). Morphology and kinetics of the three distinct phases of red blood cell invasion by Plasmodium falciparum merozoites. *Int J Parasitol* **39,** 91-96.

Glushakova, S., Humphrey, G., Leikina, E., Balaban, A., Miller, J. and Zimmerberg, J. (2010). New Stages in the Program of Malaria Parasite Egress Imaged in Normal and Sickle Erythrocytes. *Curr Biol*.

Glushakova, S., Mazar, J., Hohmann-Marriott, M.F., Hama, E. and Zimmerberg, J. (2008). Irreversible effect of cysteine protease inhibitors on the release of malaria parasites from infected erythrocytes. *Cell Microbiol*.

Glushakova, S., Yin, D., Li, T. and Zimmerberg, J. (2005). Membrane transformation during malaria parasite release from human red blood cells. *Curr Biol* **15,** 1645-1650.

Gluzman, I.Y., Francis, S.E., Oksman, A., Smith, C.E., Duffin, K.L. and Goldberg, D.E. (1994). Order and specificity of the Plasmodium falciparum hemoglobin degradation pathway. *J Clin Invest* **93,** 1602-1608.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23,** 175-182.

Goel, V.K., Li, X., Chen, H., Liu, S.C., Chishti, A.H. and Oh, S.S. (2003). Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes. *Proc Natl Acad Sci U S A* **100,** 5164-5169.

Green, J.L., Hinds, L., Grainger, M., Knuepfer, E. and Holder, A.A. (2006). Plasmodium thrombospondin related apical merozoite protein (PTRAMP) is shed from the surface of merozoites by PfSUB2 upon invasion of erythrocytes. *Mol Biochem Parasitol* **150,** 114-117.

Green, J.L., Rees-Channer, R.R., Howell, S.A., Martin, S.R., Knuepfer, E., Taylor, H.M.*, et al.* (2008). The motor complex of Plasmodium falciparum: phosphorylation by a calcium-dependent protein kinase. *J Biol Chem* **283,** 30980-30989.

Greenbaum, D.C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K.F., Engel, J.*, et al.* (2002). A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **298,** 2002-2006.

Grudkowska, M. and Zagdanska, B. (2004). Multifunctional role of plant cysteine proteinases. *Acta Biochim Pol* **51,** 609-624.

Guerra, C.A., Gikandi, P.W., Tatem, A.J., Noor, A.M., Smith, D.L., Hay, S.I. and Snow, R.W. (2008). The limits and intensity of Plasmodium falciparum transmission: implications for malaria control and elimination worldwide. *PLoS Med* **5,** e38.

Haase, S. and de Koning-Ward, T.F. (2010). New insights into protein export in malaria parasites. *Cell Microbiol* **12,** 580-587.

Hackett, F., Sajid, M., Withers-Martinez, C., Grainger, M. and Blackman, M.J. (1999). PfSUB-2: a second subtilisin-like protein in Plasmodium falciparum merozoites. *Mol Biochem Parasitol* **103,** 183-195.

Hadley, T., Aikawa, M. and Miller, L.H. (1983). Plasmodium knowlesi: studies on invasion of rhesus erythrocytes by merozoites in the presence of protease inhibitors. *Exp Parasitol* **55,** 306-311.

Hakansson, S., Charron, A.J. and Sibley, L.D. (2001). Toxoplasma evacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. *EMBO J* **20,** 3132-3144.

Han, Y.S., Thompson, J., Kafatos, F.C. and Barillas-Mury, C. (2000). Molecular interactions between Anopheles stephensi midgut cells and Plasmodium berghei: the time bomb theory of ookinete invasion of mosquitoes. *EMBO J* **19,** 6030-6040.

Hansen, G., Heitmann, A., Witt, T., Li, H., Jiang, H., Shen, X.*, et al.* (2011). Structural basis for the regulation of cysteine-protease activity by a new class of protease inhibitors in Plasmodium. *Structure* **19,** 919-929.

Hanspal, M., Dua, M., Takakuwa, Y., Chishti, A.H. and Mizuno, A. (2002). Plasmodium falciparum cysteine protease falcipain-2 cleaves erythrocyte membrane skeletal proteins at late stages of parasite development. *Blood* **100,** 1048-1054.

Harlow, E. and Lane, D. (1988) Antibodies: a laboratory manual, Cold Spring Harbor Laboratory.

Harris, P.K., Yeoh, S., Dluzewski, A.R., O'Donnell, R.A., Withers-Martinez, C., Hackett, F.*, et al.* (2005). Molecular identification of a malaria merozoite surface sheddase. *PLoS Pathog* **1,** 241-251.

Hawass, Z., Gad, Y.Z., Ismail, S., Khairat, R., Fathalla, D., Hasan, N.*, et al.* (2010). Ancestry and pathology in King Tutankhamun's family. *JAMA* **303,** 638-647.

Hay, S.I., Guerra, C.A., Tatem, A.J., Noor, A.M. and Snow, R.W. (2004). The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis* **4,** 327-336.

Hay, S.I., Okiro, E.A., Gething, P.W., Patil, A.P., Tatem, A.J., Guerra, C.A. and Snow, R.W. (2010). Estimating the global clinical burden of Plasmodium falciparum malaria in 2007. *PLoS Med* **7,** e1000290.

Heussler, V.T., Machado, J., Jr., Fernandez, P.C., Botteron, C., Chen, C.G., Pearse, M.J. and Dobbelaere, D.A. (1999). The intracellular parasite Theileria parva protects infected T cells from apoptosis. *Proc Natl Acad Sci U S A* **96,** 7312-7317.

Hippocrates (2004) Of the Epidemics, Kessinger Publishing.

Hodder, A.N., Drew, D.R., Epa, V.C., Delorenzi, M., Bourgon, R., Miller, S.K.*, et al.* (2003). Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of Plasmodium falciparum SERA5. *J Biol Chem* **278,** 48169-48177.

Hodder, A.N., Malby, R.L., Clarke, O.B., Fairlie, W.D., Colman, P.M., Crabb, B.S. and Smith, B.J. (2009). Structural Insights into the Protease-like Antigen Plasmodium falciparum SERA5 and Its Noncanonical Active-Site Serine. *J Mol Biol*.

Holder, A.A. (2009). The carboxy-terminus of merozoite surface protein 1: structure, specific antibodies and immunity to malaria. *Parasitology***,** 1-12.

Holder, A.A., Blackman, M.J., Burghaus, P.A., Chappel, J.A., Ling, I.T., McCallum-Deighton, N. and Shai, S. (1992). A malaria merozoite surface protein (MSP1)-structure, processing and function. *Mem Inst Oswaldo Cruz* **87 Suppl 3,** 37-42.

Holder, A.A. and Freeman, R.R. (1982). Biosynthesis and processing of a Plasmodium falciparum schizont antigen recognized by immune serum and a monoclonal antibody. *J Exp Med* **156,** 1528-1538.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R.*, et al.* (2002). The genome sequence of the malaria mosquito Anopheles gambiae. *Science* **298,** 129-149.

Homer and Jarnot, L. (2007) The Iliad Book XXII: The Death of Hector, BookThug.

Hooper, N.M. (2002). Proteases: a primer. *Essays Biochem* **38,** 1-8.

Horta, M.F. (1997). Pore-forming proteins in pathogenic protozoan parasites. *Trends Microbiol* **5,** 363-366.

Hu, K., Roos, D.S. and Murray, J.M. (2002). A novel polymer of tubulin forms the conoid of Toxoplasma gondii. *J Cell Biol* **156,** 1039-1050.

Hughes, G.L., Koga, R., Xue, P., Fukatsu, T. and Rasgon, J.L. (2011). Wolbachia infections are virulent and inhibit the human malaria parasite Plasmodium falciparum in anopheles gambiae. *PLoS Pathog* **7,** e1002043.

Humphreys, M.J., Moon, R.P., Klinder, A., Fowler, S.D., Rupp, K., Bur, D.*, et al.* (1999). The aspartic proteinase from the rodent parasite Plasmodium berghei as a potential model for plasmepsins from the human malaria parasite, Plasmodium falciparum. *FEBS Lett* **463,** 43-48.

Hung, S.S. and Parkinson, J. (2011). Post-genomics resources and tools for studying apicomplexan metabolism. *Trends Parasitol* **27,** 131-140.

Imwong, M., Snounou, G., Pukrittayakamee, S., Tanomsing, N., Kim, J.R., Nandy, A.*, et al.* (2007). Relapses of Plasmodium vivax infection usually result from activation of heterologous hypnozoites. *J Infect Dis* **195,** 927-933.

Inselburg, J., Bzik, D.J., Li, W.B., Green, K.M., Kansopon, J., Hahm, B.K.*, et al.* (1991). Protective immunity induced in Aotus monkeys by recombinant SERA proteins of Plasmodium falciparum. *Infect Immun* **59,** 1247-1250.

Isaacs, A.T., Li, F., Jasinskiene, N., Chen, X., Nirmala, X., Marinotti, O.*, et al.* (2011). Engineered Resistance to Plasmodium falciparum Development in Transgenic Anopheles stephensi. *PLoS Pathog* **7,** e1002017.

Janouskovec, J., Horak, A., Obornik, M., Lukes, J. and Keeling, P.J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci U S A* **107,** 10949-10954.

Janse, C.J. and Waters, A.P. (2007). The exoneme helps malaria parasites to break out of blood cells. *Cell* **131,** 1036-1038.

Jean, L., Hackett, F., Martin, S.R. and Blackman, M.J. (2003). Functional characterization of the propeptide of Plasmodium falciparum subtilisin-like protease-1. *J Biol Chem* **278,** 28572-28579.

Jones, J.L., Kruszon-Moran, D., Wilson, M., McQuillan, G., Navin, T. and McAuley, J.B. (2001). Toxoplasma gondii infection in the United States: seroprevalence and risk factors. *Am J Epidemiol* **154,** 357-365.

Joy, D.A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A.U.*, et al.* (2003). Early origin and recent expansion of Plasmodium falciparum. *Science* **300,** 318-321.

Kadekoppala, M. and Holder, A.A. (2010). Merozoite surface proteins of the malaria parasite: the MSP1 complex and the MSP7 family. *Int J Parasitol* **40,** 1155-1161.

Kadekoppala, M., O'Donnell, R.A., Grainger, M., Crabb, B.S. and Holder, A.A. (2008). Deletion of the Plasmodium falciparum merozoite surface protein 7 gene impairs parasite invasion of erythrocytes. *Eukaryot Cell* **7,** 2123-2132.

Kadota, K., Ishino, T., Matsuyama, T., Chinzei, Y. and Yuda, M. (2004). Essential role of membrane-attack protein in malarial transmission to mosquito host. *Proc Natl Acad Sci U S A* **101,** 16310-16315.

Kafsack, B.F. and Carruthers, V.B. (2010). Apicomplexan perforin-like proteins. *Commun Integr Biol* **3,** 18-23.

Kafsack, B.F., Pena, J.D., Coppens, I., Ravindran, S., Boothroyd, J.C. and Carruthers, V.B. (2008). Rapid Membrane Disruption by a Perforin-Like Protein Facilitates Parasite Exit from Host Cells. *Science*.

Kanaoka, Y., Takahashi, T., Nakayama, H., Takada, K., Kimura, T., and Sakakibara, S. (1977). Organic fluorescence reagent. IV. Synthesis of a key fluorogenic amide, Larginine-
4-methylcoumaryl-7-amide (L-Arg-MCA) and its derivatives. Fluorescence

assays for trypsin and papain. *Chem Pharm Bull* **25,** 3126-3128.
Kara, U.A., Stenzel, D.J., Ingram, L.T., Bushell, G.R., Lopez, J.A. and Kidson, C. (1988). Inhibitory monoclonal antibody against a (myristylated) small-molecular-weight antigen from Plasmodium falciparum associated with the parasitophorous vacuole membrane. *Infect Immun* **56,** 903-909.
Kauth, C.W., Woehlbier, U., Kern, M., Mekonnen, Z., Lutz, R., Mucke, N.*, et al.* (2006). Interactions between merozoite surface proteins 1, 6, and 7 of the malaria parasite Plasmodium falciparum. *J Biol Chem* **281,** 31517-31527.
Keeley, A. and Soldati, D. (2004). The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends Cell Biol* **14,** 528-532.
Keiser, J., Utzinger, J., Caldas de Castro, M., Smith, T.A., Tanner, M. and Singer, B.H. (2004). Urbanization in sub-saharan Africa and implication for malaria control. *Am J Trop Med Hyg* **71,** 118-127.
Kerr, I.D., Lee, J.H., Pandey, K.C., Harrison, A., Sajid, M., Rosenthal, P.J. and Brinen, L.S. (2009). Structures of Falcipain-2 and Falcipain-3 Bound to Small Molecule Inhibitors: Implications for Substrate Specificity (double dagger). *J Med Chem*.
Kiefer, M.C., Crawford, K.A., Boley, L.J., Landsberg, K.E., Gibson, H.L., Kaslow, D.C. and Barr, P.J. (1996). Identification and cloning of a locus of serine repeat antigen (sera)-related genes from Plasmodium vivax. *Mol Biochem Parasitol* **78,** 55-65.
Kieschnick, H., Wakefield, T., Narducci, C.A. and Beckers, C. (2001). Toxoplasma gondii attachment to host cells is regulated by a calmodulin-like domain protein kinase. *J Biol Chem* **276,** 12369-12377.
Kilejian, A. (1975). Circular mitochondrial DNA from the avian malarial parasite Plasmodium lophurae. *Biochim Biophys Acta* **390,** 276-284.
Kim, K. (2004). Role of proteases in host cell invasion by Toxoplasma gondii and other Apicomplexa. *Acta Trop* **91,** 69-81.
Kim, K. and Weiss, L.M. (2008). Toxoplasma: the next 100years. *Microbes Infect* **10,** 978-984.
Kitjaroentham, A., Suthiphongchai, T. and Wilairat, P. (2006). Effect of metalloprotease inhibitors on invasion of red blood cell by Plasmodium falciparum. *Acta Trop* **97,** 5-9.
Klayman, D.L. (1985). Qinghaosu (artemisinin): an antimalarial drug from China. *Science* **228,** 1049-1055.
Klemba, M. and Goldberg, D.E. (2002). Biological roles of proteases in parasitic protozoa. *Annu Rev Biochem* **71,** 275-305.
Klemba, M. and Goldberg, D.E. (2005). Characterization of plasmepsin V, a membrane-bound aspartic protease homolog in the endoplasmic reticulum of Plasmodium falciparum. *Mol Biochem Parasitol* **143,** 183-191.
Knapp, B., Hundt, E., Nau, U. and Kupper, H.A. (1989). Molecular cloning, genomic structure and localization in a blood stage antigen of Plasmodium falciparum characterized by a serine stretch. *Mol Biochem Parasitol* **32,** 73-83.
Knapp, B., Nau, U., Hundt, E. and Kupper, H.A. (1991). A new blood stage antigen of Plasmodium falciparum highly homologous to the serine-stretch protein SERP. *Mol Biochem Parasitol* **44,** 1-13.
Knipling, E.F. (1959). Sterile-male method of population control. *Science* **130,** 902-904.
Kooij, T.W., Carlton, J.M., Bidwell, S.L., Hall, N., Ramesar, J., Janse, C.J. and Waters, A.P. (2005). A Plasmodium whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog* **1,** e44.
Koussis, K. (2009) Thesis.
Koussis, K., Withers-Martinez, C., Yeoh, S., Child, M., Hackett, F., Knuepfer, E.*, et al.* (2009). A multifunctional serine protease primes the malaria parasite for red blood cell invasion. *EMBO J*.
Kumar, A., Kumar, K., Korde, R., Puri, S.K., Malhotra, P. and Singh Chauhan, V. (2007). Falcipain-1, a Plasmodium falciparum cysteine protease with vaccine potential. *Infect Immun* **75,** 2026-2034.
Kumar, K.A., Sano, G., Boscardin, S., Nussenzweig, R.S., Nussenzweig, M.C., Zavala, F. and Nussenzweig, V. (2006). The circumsporozoite protein is an immunodominant protective antigen in irradiated sporozoites. *Nature* **444,** 937-940.
Kyes, S., Horrocks, P. and Newbold, C. (2001). Antigenic variation at the infected red cell surface in malaria. *Annu Rev Microbiol* **55,** 673-707.
Kyes, S.A., Kraemer, S.M. and Smith, J.D. (2007). Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multigene family. *Eukaryot Cell* **6,** 1511-1520.

LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R.*, et al.* (2005). A protein interaction network of the malaria parasite Plasmodium falciparum. *Nature* **438,** 103-107.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680-685.

Lambros, C. and Vanderberg, J.P. (1979). Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J Parasitol* **65,** 418-420.

Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M., Pain, A., Sauerwein, R.W.*, et al.* (2002). Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. *Nature* **419,** 537-542.

Le Bonniec, S., Deregnaucourt, C., Redeker, V., Banerjee, R., Grellier, P., Goldberg, D.E. and Schrevel, J. (1999). Plasmepsin II, an acidic hemoglobinase from the Plasmodium falciparum food vacuole, is active at neutral pH on the host erythrocyte membrane skeleton. *J Biol Chem* **274,** 14218-14223.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M.*, et al.* (2004). Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome Res* **14,** 2308-2318.

Lebrun, M., Michelin, A., El Hajj, H., Poncet, J., Bradley, P.J., Vial, H. and Dubremetz, J.F. (2005). The rhoptry neck protein RON4 re-localizes at the moving junction during Toxoplasma gondii invasion. *Cell Microbiol* **7,** 1823-1833.

Lee, K.S., Cox-Singh, J. and Singh, B. (2009). Morphological features and differential counts of Plasmodium knowlesi parasites in naturally acquired human infections. *Malar J* **8,** 73.

Lee, K.S., Divis, P.C., Zakaria, S.K., Matusop, A., Julin, R.A., Conway, D.J.*, et al.* (2011). Plasmodium knowlesi: Reservoir Hosts and Tracking the Emergence in Humans and Macaques. *PLoS Pathog* **7,** e1002015.

Levine, N.D. (1988). Progress in taxonomy of the Apicomplexan protozoa. *J Protozool* **35,** 518-520.

Lew, V.L. (2011). Malaria: surprising mechanism of merozoite egress revealed. *Curr Biol* **21,** R314-316.

Lew, V.L. and Tiffert, T. (2007). Is invasion efficiency in malaria controlled by pre-invasion events? *Trends Parasitol* **23,** 481-484.

Li, J., Matsuoka, H., Mitamura, T. and Horii, T. (2002a). Characterization of proteases involved in the processing of Plasmodium falciparum serine repeat antigen (SERA). *Mol Biochem Parasitol* **120,** 177-186.

Li, J., Mitamura, T., Fox, B.A., Bzik, D.J. and Horii, T. (2002b). Differential localization of processed fragments of Plasmodium falciparum serine repeat antigen and further processing of its N-terminal 47 kDa fragment. *Parasitol Int* **51,** 343-352.

Li, X., Chen, H., Oo, T.H., Daly, T.M., Bergman, L.W., Liu, S.C.*, et al.* (2004). A co-ligand complex anchors Plasmodium falciparum merozoites to the erythrocyte invasion receptor band 3. *J Biol Chem* **279,** 5765-5771.

Lingelbach, K. and Joiner, K.A. (1998). The parasitophorous vacuole membrane surrounding Plasmodium and Toxoplasma: an unusual compartment in infected cells. *J Cell Sci* **111 (Pt 11),** 1467-1475.

Liu, J., Gluzman, I.Y., Drew, M.E. and Goldberg, D.E. (2005). The role of Plasmodium falciparum food vacuole plasmepsins. *J Biol Chem* **280,** 1432-1437.

Lopez, C., Saravia, C., Gomez, A., Hoebeke, J. and Patarroyo, M.A. (2010). Mechanisms of genetically-based resistance to malaria. *Gene* **467,** 1-12.

Lyon, J.A. and Haynes, J.D. (1986). Plasmodium falciparum antigens synthesized by schizonts and stabilized at the merozoite surface when schizonts mature in the presence of protease inhibitors. *J Immunol* **136,** 2245-2251.

Matthews, B.W., Sigler, P.B., Henderson, R. and Blow, D.M. (1967). Three-dimensional structure of tosyl-alpha-chymotrypsin. *Nature* **214,** 652-656.

Matthews, D.A., Alden, R.A., Birktoft, J.J., Freer, T. and Kraut, J. (1977). Re-examination of the charge relay system in subtilisin comparison with other serine proteases. *J Biol Chem* **252,** 8875-8883.

Maude, R.J., Woodrow, C.J. and White, L.J. (2010). Artemisinin Antimalarials: Preserving the "Magic Bullet". *Drug Dev Res* **71,** 12-19.

McCoubrie, J.E., Miller, S.K., Sargeant, T., Good, R.T., Hodder, A.N., Speed, T.P.*, et al.* (2007). Evidence for a common role for the serine-type Plasmodium falciparum serine repeat antigen proteases: implications for vaccine and drug design. *Infect Immun* **75,** 5565-5574.

McFadden, G.I. (2010). The apicoplast. *Protoplasma*.

McKerrow, J.H., Caffrey, C., Kelly, B., Loke, P. and Sajid, M. (2006). Proteases in parasitic diseases. *Annu Rev Pathol* **1,** 497-536.

McKerrow, J.H., Engel, J.C. and Caffrey, C.R. (1999). Cysteine protease inhibitors as chemotherapy for parasitic infections. *Bioorg Med Chem* **7,** 639-644.

McRobert, L., Taylor, C.J., Deng, W., Fivelman, Q.L., Cummings, R.M., Polley, S.D*., et al.* (2008). Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. *PLoS Biol* **6,** e139.

Menard, R. (2001). Gliding motility and cell invasion by Apicomplexa: insights from the Plasmodium sporozoite. *Cell Microbiol* **3,** 63-73.

Mendis, K., Sina, B.J., Marchesini, P. and Carter, R. (2001). The neglected burden of Plasmodium vivax malaria. *Am J Trop Med Hyg* **64,** 97-106.

Miller, L.H., Baruch, D.I., Marsh, K. and Doumbo, O.K. (2002a). The pathogenic basis of malaria. *Nature* **415,** 673-679.

Miller, R.L., Ikram, S., Armelagos, G.J., Walker, R., Harer, W.B., Shiff, C.J*., et al.* (1994). Diagnosis of Plasmodium falciparum infections in mummies using the rapid manual ParaSight-F test. *Trans R Soc Trop Med Hyg* **88,** 31-32.

Miller, S.K., Good, R.T., Drew, D.R., Delorenzi, M., Sanders, P.R., Hodder, A.N*., et al.* (2002b). A subset of Plasmodium falciparum SERA genes are expressed and appear to play an important role in the erythrocytic cycle. *J Biol Chem* **277,** 47524-47532.

Millholland, M.G., Chandramohanadas, R., Pizarro, A., Wehr, A., Shi, H., Darling, C*., et al.* (2011). The malaria parasite progressively dismantles the host erythrocyte cytoskeleton for efficient egress. *Mol Cell Proteomics*.

Mital, J., Meissner, M., Soldati, D. and Ward, G.E. (2005). Conditional expression of Toxoplasma gondii apical membrane antigen-1 (TgAMA1) demonstrates that TgAMA1 plays a critical role in host cell invasion. *Mol Biol Cell* **16,** 4341-4349.

Mitchell, G.H., Thomas, A.W., Margos, G., Dluzewski, A.R. and Bannister, L.H. (2004). Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. *Infect Immun* **72,** 154-158.

Moonen, B., Cohen, J.M., Snow, R.W., Slutsker, L., Drakeley, C., Smith, D.L*., et al.* (2010). Operational strategies to achieve and maintain malaria elimination. *Lancet* **376,** 1592-1603.

Morrison, W.I. and McKeever, D.J. (2006). Current status of vaccine development against Theileria parasites. *Parasitology* **133 Suppl,** S169-187.

Morrissette, N.S. and Sibley, L.D. (2002). Cytoskeleton of apicomplexan parasites. *Microbiol Mol Biol Rev* **66,** 21-38; table of contents.

Mueller, I., Zimmerman, P.A. and Reeder, J.C. (2007). Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. *Trends Parasitol* **23,** 278-283.

Muller, W.E., Boreiko, A., Wang, X., Belikov, S.I., Wiens, M., Grebenjuk, V.A*., et al.* (2007). Silicateins, the major biosilica forming enzymes present in demosponges: protein analysis and phylogenetic relationship. *Gene* **395,** 62-71.

Mullis, K.B. (1990). Target amplification for DNA analysis by the polymerase chain reaction. *Ann Biol Clin (Paris)* **48,** 579-582.

Murata, C.E. and Goldberg, D.E. (2003). Plasmodium falciparum falcilysin: a metalloprotease with dual specificity. *J Biol Chem* **278,** 38022-38028.

Nacher, M., Singhasivanon, P., Silachamroon, U., Treeprasertsuk, S., Tosukhowong, T., Vannaphan, S*., et al.* (2002). Decreased hemoglobin concentrations, hyperparasitemia, and severe malaria are associated with increased Plasmodium falciparum gametocyte carriage. *J Parasitol* **88,** 97-101.

Nerlich, A.G., Schraut, B., Dittrich, S., Jelinek, T. and Zink, A.R. (2008). Plasmodium falciparum in ancient Egypt. *Emerg Infect Dis* **14,** 1317-1319.

Newton, C.R. and Krishna, S. (1998a). Severe falciparum malaria in children: current understanding of pathophysiology and supportive treatment. *Pharmacol Ther* **79,** 1-53.

Newton, C.R., Taylor, T.E. and Whitten, R.O. (1998b). Pathophysiology of fatal falciparum malaria in African children. *Am J Trop Med Hyg* **58,** 673-683.

Ni, M. (1995) The Yellow Emperor's Classic of medicine: a new translation of the Neijing Suwen with commentary, Shambhala.

Nichols, B.A. and Chiappino, M.L. (1987). Cytoskeleton of Toxoplasma gondii. *J Protozool* **34,** 217-226.

Nomura, T. and Katunuma, N. (2005). Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells. *J Med Invest* **52,** 1-9.

Nyalwidhe, J., Baumeister, S., Hibbs, A.R., Tawill, S., Papakrivos, J., Volker, U. and Lingelbach, K. (2002). A nonpermeant biotin derivative gains access to the parasitophorous vacuole in

Plasmodium falciparum-infected erythrocytes permeabilized with streptolysin O. *J Biol Chem* **277,** 40005-40011.

Nyalwidhe, J. and Lingelbach, K. (2006). Proteases and chaperones are the most abundant proteins in the parasitophorous vacuole of Plasmodium falciparum-infected erythrocytes. *Proteomics* **6,** 1563-1573.

O'Donnell, R.A. and Blackman, M.J. (2005). The role of malaria merozoite proteases in red blood cell invasion. *Curr Opin Microbiol* **8,** 422-427.

O'Donnell, R.A., Hackett, F., Howell, S.A., Treeck, M., Struck, N., Krnajski, Z.*, et al.* (2006). Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. *J Cell Biol* **174,** 1023-1033.

O'Donnell, R.A., Saul, A., Cowman, A.F. and Crabb, B.S. (2000). Functional conservation of the malaria vaccine antigen MSP-119across distantly related Plasmodium species. *Nat Med* **6,** 91-95.

Ogata, K., Arakawa, M., Kasahara, T., Shioiri-Nakano, K. and Hiraoka, K. (1983). Detection of toxoplasma membrane antigens transferred from SDS-polyacrylamide gel to nitrocellulose with monoclonal antibody and avidin-biotin, peroxidase anti-peroxidase and immunoperoxidase methods. *J Immunol Methods* **65,** 75-82.

Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M. and Suzuki, K. (1984). Evolutionary origin of a calcium-dependent protease by fusion of genes for a thiol protease and a calcium-binding protein? *Nature* **312,** 566-570.

Okech, B., Mujuzi, G., Ogwal, A., Shirai, H., Horii, T. and Egwang, T.G. (2006). High titers of IgG antibodies against Plasmodium falciparum serine repeat antigen 5 (SERA5) are associated with protection against severe malaria in Ugandan children. *Am J Trop Med Hyg* **74,** 191-197.

Okech, B.A., Nalunkuma, A., Okello, D., Pang, X.L., Suzue, K., Li, J.*, et al.* (2001). Natural human immunoglobulin G subclass responses to Plasmodium falciparum serine repeat antigen in Uganda. *Am J Trop Med Hyg* **65,** 912-917.

Olaya, P. and Wasserman, M. (1991). Effect of calpain inhibitors on the invasion of human erythrocytes by the parasite Plasmodium falciparum. *Biochim Biophys Acta* **1096,** 217-221.

Omara-Opyene, A.L., Moura, P.A., Sulsona, C.R., Bonilla, J.A., Yowell, C.A., Fujioka, H.*, et al.* (2004). Genetic disruption of the Plasmodium falciparum digestive vacuole plasmepsins demonstrates their functional redundancy. *J Biol Chem* **279,** 54088-54096.

Opitz, C., Di Cristina, M., Reiss, M., Ruppert, T., Crisanti, A. and Soldati, D. (2002). Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite Toxoplasma gondii. *EMBO J* **21,** 1577-1585.

Paaijmans, K.P., Read, A.F. and Thomas, M.B. (2009). Understanding the link between malaria risk and climate. *Proc Natl Acad Sci U S A* **106,** 13844-13849.

Pachebat, J.A., Kadekoppala, M., Grainger, M., Dluzewski, A.R., Gunaratne, R.S., Scott-Finnigan, T.J.*, et al.* (2007). Extensive proteolytic processing of the malaria parasite merozoite surface protein 7 during biosynthesis and parasite release from erythrocytes. *Mol Biochem Parasitol* **151,** 59-69.

Pang, X.L. and Horii, T. (1998). Complement-mediated killing of Plasmodium falciparum erythrocytic schizont with antibodies to the recombinant serine repeat antigen (SERA). *Vaccine* **16,** 1299-1305.

Pang, X.L., Mitamura, T. and Horii, T. (1999). Antibodies reactive with the N-terminal domain of Plasmodium falciparum serine repeat antigen inhibit cell proliferation by agglutinating merozoites and schizonts. *Infect Immun* **67,** 1821-1827.

Pasteur (1881). On the Germ Theory. *Science* **2,** 420-422.

Payne, D. (1987). Spread of chloroquine resistance in Plasmodium falciparum. *Parasitol Today* **3,** 241-246.

Pei, Z. and Sebti, S.M. (1996). Cys102 and His398 are required for bleomycin-inactivating activity but not for hexamer formation of yeast bleomycin hydrolase. *Biochemistry* **35,** 10751-10756.

Perrin, L.H. and Dayal, R. (1982). Immunity to asexual erythrocytic stages of Plasmodium falciparum: role of defined antigens in the humoral response. *Immunol Rev* **61,** 245-269.

Perrin, L.H., Merkli, B., Gabra, M.S., Stocker, J.W., Chizzolini, C. and Richle, R. (1985). Immunization with a Plasmodium falciparum merozoite surface antigen induces a partial immunity in monkeys. *J Clin Invest* **75,** 1718-1721.

Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. and Richle, R. (1984). Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages. *J Exp Med* **160,** 441-451.

Perrin, L.H., Ramirez, E., Lambert, P.H. and Miescher, P.A. (1981). Inhibition of P. falciparum growth in human erythrocytes by monoclonal antibodies. *Nature* **289,** 301-303.

Peters, J.M., Franke, W.W. and Kleinschmidt, J.A. (1994). Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm. *J Biol Chem* **269,** 7709-7718.

Pierce, J.A., Robertson, C.R. and Leighton, T.J. (1992). Physiological and genetic strategies for enhanced subtilisin production by Bacillus subtilis. *Biotechnol Prog* **8,** 211-218.

Podgorski, I. and Sloane, B.F. (2003). Cathepsin B and its role(s) in cancer progression. *Biochem Soc Symp***,** 263-276.

Polgar, L. (2005). The catalytic triad of serine peptidases. *Cell Mol Life Sci* **62,** 2161-2172.

Polonais, V. and Soldati-Favre, D. (2010). Versatility in the acquisition of energy and carbon sources by the Apicomplexa. *Biol Cell* **102,** 435-445.

Ponpuak, M., Klemba, M., Park, M., Gluzman, I.Y., Lamppa, G.K. and Goldberg, D.E. (2007). A role for falcilysin in transit peptide degradation in the Plasmodium falciparum apicoplast. *Mol Microbiol* **63,** 314-334.

Potgieter, F.T. and Els, H.J. (1977). The fine structure of intra-erythrocytic stages of Babesia bigemina. *Onderstepoort J Vet Res* **44,** 157-168.

Preiser, P., Kaviratne, M., Khan, S., Bannister, L. and Jarra, W. (2000). The apical organelles of malaria merozoites: host cell selection, invasion, host immunity and immune evasion. *Microbes Infect* **2,** 1461-1477.

Preiser, P.R., Khan, S., Costa, F.T., Jarra, W., Belnoue, E., Ogun, S.*, et al.* (2002). Stage-specific transcription of distinct repertoires of a multigene family during Plasmodium life cycle. *Science* **295,** 342-345.

Price, R.N., Uhlemann, A.C., Brockman, A., McGready, R., Ashley, E., Phaipun, L.*, et al.* (2004). Mefloquine resistance in Plasmodium falciparum and increased pfmdr1 gene copy number. *Lancet* **364,** 438-447.

Prudencio, M. and Mota, M.M. (2007). To migrate or to invade: those are the options. *Cell Host Microbe* **2,** 286-288.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E.*, et al.* (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24,** 218-229.

Putrianti, E.D., Schmidt-Christensen, A., Arnold, I., Heussler, V.T., Matuschewski, K. and Silvie, O. (2009). The Plasmodium serine-type SERA proteases display distinct expression patterns and non-essential in vivo roles during life cycle progression of the malaria parasite. *Cell Microbiol*.

Ramasamy, R., Ramasamy, M. and Yasawardena, S. (2001). Antibodies and Plasmodium falciparum merozoites. *Trends Parasitol* **17,** 194-197.

Ranson, H., Jensen, B., Wang, X., Prapanthadara, L., Hemingway, J. and Collins, F.H. (2000). Genetic mapping of two loci affecting DDT resistance in the malaria vector Anopheles gambiae. *Insect Mol Biol* **9,** 499-507.

Rape, M. and Jentsch, S. (2002). Taking a bite: proteasomal protein processing. *Nature cell biology* **4,** E113-116.

Rawlings, N.D. and Barrett, A.J. (1993). Evolutionary families of peptidases. *Biochem J* **290 (Pt 1),** 205-218.

Rawlings, N.D. and Barrett, A.J. (1994). Families of serine peptidases. *Methods Enzymol* **244,** 19-61.

Rawlings, N.D., Barrett, A.J. and Bateman, A. (2010). MEROPS: the peptidase database. *Nucleic Acids Res* **38,** D227-233.

Rayner, J.C., Vargas-Serrato, E., Huber, C.S., Galinski, M.R. and Barnwell, J.W. (2001). A Plasmodium falciparum homologue of Plasmodium vivax reticulocyte binding protein (PvRBP1) defines a trypsin-resistant erythrocyte invasion pathway. *J Exp Med* **194,** 1571-1581.

Rennenberg, A., Lehmann, C., Heitmann, A., Witt, T., Hansen, G., Nagarajan, K.*, et al.* (2010). Exoerythrocytic Plasmodium parasites secrete a cysteine protease inhibitor involved in sporozoite invasion and capable of blocking cell death of host hepatocytes. *PLoS Pathog* **6,** e1000825.

Reyburn, H., Mbatia, R., Drakeley, C., Bruce, J., Carneiro, I., Olomi, R.*, et al.* (2005). Association of transmission intensity and age with clinical manifestations and case fatality of severe Plasmodium falciparum malaria. *JAMA* **293,** 1461-1470.

Rich, S.M., Leendertz, F.H., Xu, G., LeBreton, M., Djoko, C.F., Aminake, M.N.*, et al.* (2009). The origin of malignant malaria. *Proc Natl Acad Sci U S A* **106,** 14902-14907.

Rich, S.M., Licht, M.C., Hudson, R.R. and Ayala, F.J. (1998). Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of Plasmodium falciparum. *Proc Natl Acad Sci U S A* **95,** 4425-4430.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17,** 1030-1032.

Riley, E.M., Wahl, S., Perkins, D.J. and Schofield, L. (2006). Regulating immunity to malaria. *Parasite Immunol* **28,** 35-49.

Rivadeneira, E.M., Wasserman, M. and Espinal, C.T. (1983). Separation and concentration of schizonts of Plasmodium falciparum by Percoll gradients. *J Protozool* **30,** 367-370.

Robert, V., Macintyre, K., Keating, J., Trape, J.F., Duchemin, J.B., Warren, M. and Beier, J.C. (2003). Malaria transmission in urban sub-Saharan Africa. *Am J Trop Med Hyg* **68,** 169-176.

Roberts, D.J., Craig, A.G., Berendt, A.R., Pinches, R., Nash, G., Marsh, K. and Newbold, C.I. (1992). Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357,** 689-692.

Rosenthal, P.J. (2002). Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr Opin Hematol* **9,** 140-145.

Rosenthal, P.J. (2004). Cysteine proteases of malaria parasites. *Int J Parasitol* **34,** 1489-1499.

Rosenthal, P.J., Kim, K., McKerrow, J.H. and Leech, J.H. (1987). Identification of three stage-specific proteinases of Plasmodium falciparum. *J Exp Med* **166,** 816-821.

Rosenthal, P.J., McKerrow, J.H., Aikawa, M., Nagasawa, H. and Leech, J.H. (1988). A malarial cysteine proteinase is necessary for hemoglobin degradation by Plasmodium falciparum. *J Clin Invest* **82,** 1560-1566.

Rosenthal, P.J. and Nelson, R.G. (1992). Isolation and characterization of a cysteine proteinase gene of Plasmodium falciparum. *Mol Biochem Parasitol* **51,** 143-152.

Ross, R. (1897). On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood. *Br Med J* **2,** 1786-1788.

Russell, P.F. (1955) Man's mastery of malaria, Oxford University Press.

Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A. and Goldberg, D.E. (2009a). Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. *Nature* **463,** 632-636.

Russo, I., Oksman, A., Vaupel, B. and Goldberg, D.E. (2009b). A calpain unique to alveolates is essential in Plasmodium falciparum and its knockdown reveals an involvement in pre-S-phase development. *Proc Natl Acad Sci U S A*.

Sabnis, Y.A., Desai, P.V., Rosenthal, P.J. and Avery, M.A. (2003). Probing the structure of falcipain-3, a cysteine protease from Plasmodium falciparum: comparative protein modeling and docking studies. *Protein Sci* **12,** 501-509.

Sachs, J. and Malaney, P. (2002). The economic and social burden of malaria. *Nature* **415,** 680-685.

Sajid, M., Withers-Martinez, C. and Blackman, M.J. (2000). Maturation and specificity of Plasmodium falciparum subtilisin-like protease-1, a malaria merozoite subtilisin-like serine protease. *J Biol Chem* **275,** 631-641.

Salmon, B.L., Oksman, A. and Goldberg, D.E. (2001). Malaria parasite exit from the host erythrocyte: a two-step process requiring extraerythrocytic proteolysis. *Proc Natl Acad Sci U S A* **98,** 271-276.

Sambrook J., G. (1989). Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*.

Sanders, P.R., Gilson, P.R., Cantin, G.T., Greenbaum, D.C., Nebl, T., Carucci, D.J.*, et al.* (2005). Distinct protein classes including novel merozoite surface antigens in Raft-like membranes of Plasmodium falciparum. *J Biol Chem* **280,** 40169-40176.

Santa-Olalla Peralta, P., Vazquez-Torres, M.C., Latorre-Fandos, E., Mairal-Claver, P., Cortina-Solano, P., Puy-Azon, A.*, et al.* (2010). First autochthonous malaria case due to Plasmodium vivax since eradication, Spain, October 2010. *Euro Surveill* **15,** 19684.

Santos, J.M., Ferguson, D.J., Blackman, M.J. and Soldati-Favre, D. (2011). Intramembrane cleavage of AMA1 triggers Toxoplasma to switch from an invasive to a replicative mode. *Science* **331,** 473-477.

Saralamba, S., Pan-Ngum, W., Maude, R.J., Lee, S.J., Tarning, J., Lindegardh, N.*, et al.* (2011). Intrahost modeling of artemisinin resistance in Plasmodium falciparum. *Proc Natl Acad Sci U S A* **108,** 397-402.

Schechter, I. and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27,** 157-162.

Schmidt-Christensen, A., Sturm, A., Horstmann, S. and Heussler, V.T. (2008). Expression and processing of Plasmodium berghei SERA3 during liver stages. *Cell Microbiol* **10,** 1723-1734.

Shaw, M.K. (1997). The same but different: the biology of Theileria sporozoite entry into bovine cells. *Int J Parasitol* **27,** 457-474.

Shea, M., Jakle, U., Liu, Q., Berry, C., Joiner, K.A. and Soldati-Favre, D. (2007). A family of aspartic proteases and a novel, dynamic and cell-cycle-dependent protease localization in the secretory pathway of Toxoplasma gondii. *Traffic* **8,** 1018-1034.

Shenai, B.R. and Rosenthal, P.J. (2002). Reducing requirements for hemoglobin hydrolysis by Plasmodium falciparum cysteine proteases. *Mol Biochem Parasitol* **122,** 99-104.

Shenai, B.R., Sijwali, P.S., Singh, A. and Rosenthal, P.J. (2000). Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of Plasmodium falciparum. *J Biol Chem* **275,** 29000-29010.

Sherman, I.W. (1998) Malaria: parasite biology, pathogenesis, and protection, ASM Press.

Sherman, I.W., Eda, S. and Winograd, E. (2004). Erythrocyte aging and malaria. *Cell Mol Biol (Noisy-le-grand)* **50,** 159-169.

Shimizu, K., Cha, J., Stucky, G.D. and Morse, D.E. (1998). Silicatein alpha: cathepsin L-like protein in sponge biosilica. *Proc Natl Acad Sci U S A* **95,** 6234-6238.

Shirley, M.W., Smith, A.L. and Blake, D.P. (2007). Challenges in the successful control of the avian coccidia. *Vaccine* **25,** 5540-5547.

Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S., Case, S.E., Yamaga, K.M.*, et al.* (1987). Merozoite surface coat precursor protein completely protects Aotus monkeys against Plasmodium falciparum malaria. *Proc Natl Acad Sci U S A* **84,** 3014-3018.

Siezen, R.J. and Leunissen, J.A. (1997). Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci* **6,** 501-523.

Sijwali, P.S., Brinen, L.S. and Rosenthal, P.J. (2001a). Systematic optimization of expression and refolding of the Plasmodium falciparum cysteine protease falcipain-2. *Protein Expr Purif* **22,** 128-134.

Sijwali, P.S., Kato, K., Seydel, K.B., Gut, J., Lehman, J., Klemba, M.*, et al.* (2004a). Plasmodium falciparum cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proc Natl Acad Sci U S A* **101,** 8721-8726.

Sijwali, P.S. and Rosenthal, P.J. (2004b). Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum. *Proc Natl Acad Sci U S A* **101,** 4384-4389.

Sijwali, P.S., Shenai, B.R., Gut, J., Singh, A. and Rosenthal, P.J. (2001b). Expression and characterization of the Plasmodium falciparum haemoglobinase falcipain-3. *Biochem J* **360,** 481-489.

Silmon de Monerri, N.C. (2010) Exiting the Erythrocyte: Functional and Temporal Analysis of a Malarial Subtilase. In *Infection and Immunity*. London, University College London.

Silmon de Monerri, N.C., Flynn, H.R., Campos, M.G., Hackett, F., Koussis, K., Withers-Martinez, C.*, et al.* (2011). Global identification of multiple substrates for Plasmodium falciparum SUB1, an essential malarial processing protease. *Infect Immun* **79,** 1086-1097.

Silvie, O., Franetich, J.F., Charrin, S., Mueller, M.S., Siau, A., Bodescot, M.*, et al.* (2004). A role for apical membrane antigen 1 during invasion of hepatocytes by Plasmodium falciparum sporozoites. *J Biol Chem* **279,** 9490-9496.

Simmons, D., Woollett, G., Bergin-Cartwright, M., Kay, D. and Scaife, J. (1987). A malaria protein exported into a new compartment within the host erythrocyte. *Embo J* **6,** 485-491.

Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S.S., Cox-Singh, J.*, et al.* (2004). A large focus of naturally acquired Plasmodium knowlesi infections in human beings. *Lancet* **363,** 1017-1024.

Singh, N., Sijwali, P.S., Pandey, K.C. and Rosenthal, P.J. (2006). Plasmodium falciparum: biochemical characterization of the cysteine protease falcipain-2'. *Exp Parasitol* **112,** 187-192.

Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y. and Hay, S.I. (2005). The global distribution of clinical episodes of Plasmodium falciparum malaria. *Nature* **434,** 214-217.

Snow, R.W., Korenromp, E.L. and Gouws, E. (2004). Pediatric mortality in Africa: plasmodium falciparum malaria as a cause or risk? *Am J Trop Med Hyg* **71,** 16-24.

Snow, R.W., Trape, J.F. and Marsh, K. (2001). The past, present and future of childhood malaria mortality in Africa. *Trends Parasitol* **17,** 593-597.

Soldati, D. and Boothroyd, J.C. (1993). Transient transfection and expression in the obligate intracellular parasite Toxoplasma gondii. *Science* **260,** 349-352.

Soni, S., Dhawan, S., Rosen, K.M., Chafel, M., Chishti, A.H. and Hanspal, M. (2005). Characterization of events preceding the release of malaria parasite from the host red blood cell. *Blood Cells Mol Dis* **35,** 201-211.

Spaccapelo, R., Janse, C.J., Caterbi, S., Franke-Fayard, B., Bonilla, J.A., Syphard, L.M.*, et al.* (2010). Plasmepsin 4-deficient Plasmodium berghei are virulence attenuated and induce protective immunity against experimental malaria. *Am J Pathol* **176,** 205-217.

Spielmann, T., Gardiner, D.L., Beck, H.P., Trenholme, K.R. and Kemp, D.J. (2006). Organization of ETRAMPs and EXP-1 at the parasite-host cell interface of malaria parasites. *Mol Microbiol* **59,** 779-794.

Srinivasan, P., Coppens, I. and Jacobs-Lorena, M. (2009). Distinct roles of Plasmodium rhomboid 1 in parasite development and malaria pathogenesis. *PLoS Pathog* **5,** e1000262.

Stafford, W.H., Blackman, M.J., Harris, A., Shai, S., Grainger, M. and Holder, A.A. (1994). N-terminal amino acid sequence of the Plasmodium falciparum merozoite surface protein-1 polypeptides. *Mol Biochem Parasitol* **66,** 157-160.

Stafford, W.H., Gunder, B., Harris, A., Heidrich, H.G., Holder, A.A. and Blackman, M.J. (1996). A 22 kDa protein associated with the Plasmodium falciparum merozoite surface protein-1 complex. *Mol Biochem Parasitol* **80,** 159-169.

Stubbs, J., Simpson, K.M., Triglia, T., Plouffe, D., Tonkin, C.J., Duraisingh, M.T.*, et al.* (2005). Molecular mechanism for switching of P. falciparum invasion pathways into human erythrocytes. *Science* **309,** 1384-1387.

Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A.*, et al.* (2006). Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313,** 1287-1290.

Sturm, A. and Heussler, V. (2007). Live and let die: manipulation of host hepatocytes by exoerythrocytic Plasmodium parasites. *Med Microbiol Immunol* **196,** 127-133.

Subbian, E., Yabuta, Y. and Shinde, U.P. (2005). Folding pathway mediated by an intramolecular chaperone: intrinsically unstructured propeptide modulates stochastic activation of subtilisin. *J Mol Biol* **347,** 367-383.

Sutherland, C.J., Laundy, M., Price, N., Burke, M., Fivelman, Q.L., Pasvol, G.*, et al.* (2008). Mutations in the Plasmodium falciparum cytochrome b gene are associated with delayed parasite recrudescence in malaria patients treated with atovaquone-proguanil. *Malar J* **7,** 240.

Szecsi, P.B. (1992). The aspartic proteases. *Scand J Clin Lab Invest Suppl* **210,** 5-22.

Takebe, S., Witola, W.H., Schimanski, B., Gunzl, A. and Ben Mamoun, C. (2007). Purification of components of the translation elongation factor complex of Plasmodium falciparum by tandem affinity purification. *Eukaryot Cell* **6,** 584-591.

Tallant, C., Marrero, A. and Gomis-Ruth, F.X. (2010). Matrix metalloproteinases: fold and function of their catalytic domains. *Biochim Biophys Acta* **1803,** 20-28.

Teixeira, C., Gomes, J.R. and Gomes, P. (2011). Falcipains, Plasmodium falciparum cysteine proteases as key drug targets against malaria. *Curr Med Chem* **18,** 1555-1572.

Thailayil, J., Magnusson, K., Godfray, H.C., Crisanti, A. and Catteruccia, F. (2011). Spermless males elicit large-scale female responses to mating in the malaria mosquito Anopheles gambiae. *Proc Natl Acad Sci U S A* **108,** 13677-13681.

Tham, W.H., Wilson, D.W., Lopaticki, S., Schmidt, C.Q., Tetteh-Quarcoo, P.B., Barlow, P.N.*, et al.* (2010). Complement receptor 1 is the host erythrocyte receptor for Plasmodium falciparum PfRh4 invasion ligand. *Proc Natl Acad Sci U S A* **107,** 17327-17332.

Tilley, L., McFadden, G., Cowman, A. and Klonis, N. (2007). Illuminating Plasmodium falciparum-infected red blood cells. *Trends Parasitol* **23,** 268-277.

Timmann, C. and Meyer, C.G. (2010). Malaria, mummies, mutations: Tutankhamun's archaeological autopsy. *Trop Med Int Health* **15,** 1278-1280.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76,** 4350-4354.

Trager, W. (1956). The intracellular position of malarial parasites. *Trans R Soc Trop Med Hyg* **50,** 419-420.

Trager, W. (1994). Cultivation of malaria parasites. *Methods Cell Biol* **45,** 7-26.

Trager, W. and Jensen, J.B. (1976). Human malaria parasites in continuous culture. *Science* **193,** 673-675.

Triglia, T., Healer, J., Caruana, S.R., Hodder, A.N., Anders, R.F., Crabb, B.S. and Cowman, A.F. (2000). Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species. *Mol Microbiol* **38,** 706-718.

Tuteja, R. (2007). Malaria - an overview. *FEBS J* **274,** 4670-4679.

Uellner, R., Zvelebil, M.J., Hopkins, J., Jones, J., MacDougall, L.K., Morgan, B.P.*, et al.* (1997). Perforin is activated by a proteolytic cleavage during biosynthesis which reveals a phospholipid-binding C2 domain. *Embo J* **16,** 7287-7296.

Ulhoa, C.J., Sankievicz, D., Limeira, P.S. and Peberdy, J.F. (2001). Effect of tunicamycin on N-acetyl-beta-D-glucosaminidase produced by Trichoderma harzianum. *Biochim Biophys Acta* **1528,** 39-42.

Urban, S. and Freeman, M. (2002a). Intramembrane proteolysis controls diverse signalling pathways throughout evolution. *Curr Opin Genet Dev* **12,** 512-518.

Urban, S. and Freeman, M. (2003). Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol Cell* **11,** 1425-1434.

Urban, S., Lee, J.R. and Freeman, M. (2001). Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* **107,** 173-182.

Urban, S., Schlieper, D. and Freeman, M. (2002b). Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids. *Curr Biol* **12,** 1507-1512.

Uzureau, P., Barale, J.C., Janse, C.J., Waters, A.P. and Breton, C.B. (2004). Gene targeting demonstrates that the Plasmodium berghei subtilisin PbSUB2 is essential for red cell invasion and reveals spontaneous genetic recombination events. *Cell Microbiol* **6,** 65-78.

Varughese, K.I., Ahmed, F.R., Carey, P.R., Hasnain, S., Huber, C.P. and Storer, A.C. (1989). Crystal structure of a papain-E-64 complex. *Biochemistry* **28,** 1330-1332.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G.*, et al.* (2001). The sequence of the human genome. *Science* **291,** 1304-1351.

Vlachou, D., Schlegelmilch, T., Runn, E., Mendes, A. and Kafatos, F.C. (2006). The developmental migration of Plasmodium in mosquitoes. *Curr Opin Genet Dev* **16,** 384-391.

Voet, D. and Voet, J.G. (2004) Biochemistry, J. Wiley & Sons.

Voges, D., Zwickl, P. and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem* **68,** 1015-1068.

Volkman, S.K., Barry, A.E., Lyons, E.J., Nielsen, K.M., Thomas, S.M., Choi, M.*, et al.* (2001). Recent origin of Plasmodium falciparum from a single progenitor. *Science* **293,** 482-484.

Waller, R.F. and McFadden, G.I. (2005). The apicoplast: a review of the derived plastid of apicomplexan parasites. *Curr Issues Mol Biol* **7,** 57-79.

Wang, F., Krai, P., Deu, E., Bibb, B., Lauritzen, C., Pedersen, J.*, et al.* (2011). Biochemical characterization of Plasmodium falciparum dipeptidyl aminopeptidase 1. *Mol Biochem Parasitol* **175,** 10-20.

Wellems, T.E., Hayton, K. and Fairhurst, R.M. (2009). The impact of malaria parasitism: from corpuscles to communities. *J Clin Invest* **119,** 2496-2505.

Wellems, T.E. and Plowe, C.V. (2001). Chloroquine-resistant malaria. *J Infect Dis* **184,** 770-776.

Wertheimer, S.P. and Barnwell, J.W. (1989). Plasmodium vivax interaction with the human Duffy blood group glycoprotein: identification of a parasite receptor-like protein. *Exp Parasitol* **69,** 340-350.

Whitford, D. (2005) Proteins : structure and function Hoboken, NJ, John Wiley & Sons, pp. xiv, 528 p.

WHO (1999) The World Health Report.

WHO (2003) The Africa Malaria Report World Health Organization.

WHO (2008) World Malaria Report 2008. World Health Organization.

WHO (2009) World Malaria Report 2009. World Health Organisation.

WHO (2010) World Malaria Report 2010. World Health Organization.

Wickham, M.E., Culvenor, J.G. and Cowman, A.F. (2003). Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J Biol Chem* **278,** 37658-37663.

Williams, T.N., Mwangi, T.W., Roberts, D.J., Alexander, N.D., Weatherall, D.J., Wambua, S.*, et al.* (2005). An immune basis for malaria protection by the sickle cell trait. *PLoS Med* **2,** e128.

Winograd, E., Clavijo, C.A., Bustamante, L.Y. and Jaramillo, M. (1999). Release of merozoites from Plasmodium falciparum-infected erythrocytes could be mediated by a non-explosive event. *Parasitol Res* **85,** 621-624.

Withers-Martinez, C., Jean, L. and Blackman, M.J. (2004). Subtilisin-like proteases of the malaria parasite. *Mol Microbiol* **53,** 55-63.

Withers-Martinez, C., Saldanha, J.W., Ely, B., Hackett, F., O'Connor, T. and Blackman, M.J. (2002). Expression of recombinant Plasmodium falciparum subtilisin-like protease-1 in insect cells. Characterization, comparison with the parasite protease, and homology modeling. *J Biol Chem* **277,** 29698-29709.

Woehlbier, U., Epp, C., Hackett, F., Blackman, M.J. and Bujard, H. (2010). Antibodies against multiple merozoite surface antigens of the human malaria parasite Plasmodium falciparum inhibit parasite maturation and red blood cell invasion. *Malar J* **9,** 77.

Wongsrichanalai, C., Pickard, A.L., Wernsdorfer, W.H. and Meshnick, S.R. (2002). Epidemiology of drug-resistant malaria. *Lancet Infect Dis* **2,** 209-218.

Wu, Y., Wang, X., Liu, X. and Wang, Y. (2003). Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. *Genome Res* **13,** 601-616.

Xie, D., Gulnik, S., Collins, L., Gustchina, E., Suvorov, L. and Erickson, J.W. (1997). Dissection of the pH dependence of inhibitor binding energetics for an aspartic protease: direct measurement of the protonation states of the catalytic aspartic acid residues. *Biochemistry* **36,** 16166-16172.

Yabuta, Y., Subbian, E., Oiry, C. and Shinde, U. (2003). Folding pathway mediated by an intramolecular chaperone. A functional peptide chaperone designed using sequence databases. *J Biol Chem* **278,** 15246-15251.

Yeoh, S., O'Donnell, R.A., Koussis, K., Dluzewski, A.R., Ansell, K.H., Osborne, S.A*., et al.* (2007). Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* **131,** 1072-1083.

Zhou, X.W., Blackman, M.J., Howell, S.A. and Carruthers, V.B. (2004). Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. *Mol Cell Proteomics* **3,** 565-576.